

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2001, 18:18:09 ; Search time 7124 Seconds  
(without alignments)  
11995.973 Million cell updates/sec

Title: US-09-515-806-1  
Perfect score: 5525  
Sequence: 1 tcgccacagcgtccgcacc.....aatgttttatactacgtga 5525

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0  
Searched: 1344157 seqs, 7733874588 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba1.\*  
2: gb\_ba2.\*  
3: gb\_ba3.\*  
4: gb\_in1.\*  
5: gb\_in2.\*  
6: gb\_in3.\*  
7: gb\_om.\*  
8: gb\_ov.\*  
9: gb\_pat1.\*  
10: gb\_pat2.\*  
11: gb\_ph.\*  
12: gb\_pl1.\*  
13: gb\_pl2.\*  
14: gb\_pl3.\*  
15: gb\_pl4.\*  
16: em\_ba1.\*  
17: em\_ba2.\*  
18: em\_fun.\*  
19: em\_htgo\_hum.\*  
20: em\_htgo\_inv.\*  
21: em\_htgo\_rod.\*  
22: em\_htgo\_hum1.\*  
23: em\_htgo\_hum2.\*  
24: em\_htgo\_hum3.\*  
25: em\_htgo\_hum4.\*  
26: em\_htgo\_hum5.\*  
27: em\_htgo\_hum6.\*  
28: em\_htgo\_hum7.\*  
29: em\_htgo\_hum8.\*  
30: em\_htg\_inv1.\*  
31: em\_htg\_inv2.\*  
32: em\_htg\_other.\*  
33: em\_htg\_rod.\*  
34: em\_hum1.\*  
35: em\_hum2.\*  
36: em\_hum3.\*  
37: em\_hum4.\*  
38: em\_hum5.\*  
39: em\_hum6.\*  
40: em\_hum7.\*  
41: em\_in.\*  
42: em\_om.\*  
43: em\_or.\*

44: em\_ov.\*  
45: em\_pat.\*  
46: em\_ph.\*  
47: em\_pl.\*  
48: em\_ro.\*  
49: em\_sts.\*  
50: em\_sy.\*  
51: em\_un.\*  
52: em\_vi.\*  
53: gb\_sts1.\*  
54: gb\_sts2.\*  
55: gb\_sts3.\*  
56: gb\_sy.\*  
57: gb\_un.\*  
58: gb\_vil.\*  
59: gb\_vil2.\*  
60: gb\_htg1.\*  
61: gb\_htg2.\*  
62: gb\_htg3.\*  
63: gb\_htg4.\*  
64: gb\_htg5.\*  
65: gb\_htg6.\*  
66: gb\_htg7.\*  
67: gb\_htg8.\*  
68: gb\_htg9.\*  
69: gb\_htg10.\*  
70: gb\_htg11.\*  
71: gb\_htg12.\*  
72: gb\_htg13.\*  
73: gb\_htg14.\*  
74: gb\_htg15.\*  
75: gb\_htg16.\*  
76: gb\_htg17.\*  
77: gb\_htg18.\*  
78: gb\_htg19.\*  
79: gb\_htg20.\*  
80: gb\_htg21.\*  
81: gb\_htg22.\*  
82: gb\_htg23.\*  
83: gb\_htg24.\*  
84: gb\_htg25.\*  
85: gb\_pr1.\*  
86: gb\_pr2.\*  
87: gb\_pr3.\*  
88: gb\_pr4.\*  
89: gb\_pr5.\*  
90: gb\_pr6.\*  
91: gb\_pr7.\*  
92: gb\_pr8.\*  
93: gb\_pr9.\*  
94: gb\_rol.\*  
95: gb\_ro2.\*  
96: gb\_in4.\*  
97: gb\_pr10.\*  
98: em\_ba3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	5107	92.4	5163	9	AX056426 Sequence
2	4878	88.3	4994	85	AB037759 Homo sapi
3	3332	60.3	3351	93	AL157497 Homo sapi
4	2650	48.0	2986	89	AK027011 Homo sapi
5	2111	38.2	2162	93	AK027011 Homo sapi
6	1032	18.7	1138	93	AL137627 Homo sapi
7	919	16.6	1917	93	AL137676 Homo sapi
8	539	9.8	164297	62	AC012138 Homo sapi



QY 1189 acatttttagtgagcacatttagtgggtctctcttctgtgcacacctgagccactcaggcc 1248  
 Db 1146 ACATTTTAGTGAGGACATAGTGGGTCCTCTGTGCTGCACACTGAGCCACTCAGGCC 1205  
 QY 1249 ceatccctgtgcatcagcttcgcaggttacacagctcagctcgttcagggccttgatttc 1308  
 Db 1206 CCATCCCTGTGCATCAGCTTCGCAGGTACACAGCTCAGCTCCTGTGCAGGCTTGATTATC 1265  
 QY 1309 tqcacagaattctgtgtgcataaggtcctgagtgcatctaatgtctgtgtgagtcag 1368  
 Db 1266 TCCACAGCAATTCCTGTGTGATAGTTCCTGTGAGTGCATCAATTCCTGTGTGATGCAG 1325  
 QY 1369 aaggcacctgcaagattacgactatagcattctctaaagcctgcgcagacatttgcagg 1428  
 Db 1326 AAGGCACCGTCAAGATTACGACNATACATTTCTAAGCGCTCGCAGACATTTGCAAGG 1385  
 QY 1429 aggtgtgtttgagcaaacccaggtctcttttagtgacaatgctctgctctataaaacgg 1488  
 Db 1386 AGGATGTGTTGAGCAAAACCCAGTTCGTTTATGTGACAATGCTGCTGCTTATAAACGG 1445  
 QY 1489 ggaagaaaggagatgtttggcgcttggcctctctgtctgtctcctcagcacaagacagg 1548  
 Db 1446 GGAAGAAAGGAGATTTTGGCGCTTGGCCCTTCTGCTGCTCCTCAGCCAAAGGACAGG 1505  
 QY 1549 aatgtggagagtaccctgtgacctccctagtgacttaccagctgactttcaagatttc 1608  
 Db 1506 AATGTGGAGATGACCTGTGACCATCCCTAGTGACTTACCAAGCTGACTTTCAAGATTTTC 1565  
 QY 1609 taaagaaatgtgtgtgtggtgatacaaggaaagatggagtcctccagcagttgttgaac 1668  
 Db 1566 TAAAGAAATGTGTGTGATGACAAAGGAAAGATGAGTCCCGCAGCAGTTGTTGAAC 1625  
 QY 1669 acagctttataatcccccagcaaaaatgctctagtggaacaaagtcctgaagattctg 1728  
 Db 1626 ACAGCTTTATAATCCCGCAAAATGCTCTAGTGTGAACAAAGTCCCTGAAGATTCTG 1685  
 QY 1729 gaggcaagattattgagactgttactctagcaacccggtaccagtgctcctctct 1788  
 Db 1686 GAGGACAGATTATGTTGAGACTGTTATTCCTAGCAACCGCTACCCAGTGTCTGCTTCT 1745  
 QY 1789 ttatgtgagacagagacagttttcccgatacttctcattgattgttgaagaattacaattc 1848  
 Db 1746 TTAGTGTGACACACAGACAGATTTTCCCGATACTTCATTGAGTTTGAAGAAATTACAAC 1805  
 QY 1849 ttgttaaggagctttttagctgtcatcaagggtgcagaaacagttggagcgtctgct 1908  
 Db 1806 TTGGTAAAGGAGCTTTTGGAGCTGTTCATCAAGGTGCAGAAAGTTGGAGCGCTGCTGCT 1865  
 QY 1909 acgcagtgaaagcgtatcccatcaaccggccagccggcagttccgcaggtatcaagggcg 1968  
 Db 1866 AGCGAGTGAAGCGCATCCCATCAACCCGGCCAGCGCGCAGTTCCGACAGATCAAGGGCG 1925  
 QY 1969 aagtacactgctgtcacggtgcacctgagaaacattgtcgctactacaacgcttggga 2028  
 Db 1926 AAGTGACACTGTGTACGGGTGCACCATGACCATGAGAACATTTGCGCTACTTACAACGCTTGA 1985  
 QY 2029 tcgagcgacagagccgagcgagggagcgagcgagcgagcgagcgagcgagcgagcgagcg 2088  
 Db 1986 TCGAGCGCACAGGGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 2045  
 QY 2089 ccaagatgacagcgagctgcagcgggcgagcgagcgagcgagcgagcgagcgagcgagcgag 2148  
 Db 2046 CCAAGATGACCGAGTGCACCGGGCAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 2105  
 QY 2149 tagagccgc 2208  
 Db 2106 TAGAGCG 2165  
 QY 2209 aqcgctcgagcgagtcctgttcccccagccagcccgagcgagcgagcgagcgagcgagcgag 2268  
 Db 2166 AGCGCTGGCGAGTGGCGGTTTCCCGCGCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2225  
 QY 2269 acgacgaggaagcagcggtggcgcttctctccagtcctctcctgcttctcagattctg 2328

Db 2226 ACAGGAGGAGGAGCAGCGGTGGCGTCTTCTCCAGTCTCTCTCCGCTGCTTCAGATCTCG 2285  
 QY 2329 aaagtgtatattatttttgacaatgaagatgaagaaagtcagaaatcagaaatcagaaatcag 2388  
 Db 2286 AAAGTGATATTATCTTTGACAATGAAGATGAGAACAGTAAAGTCAGAATCAGGATGAAG 2345  
 QY 2389 attgcaatgaagaatgctgcatgaagaatgaagaatgaagaatgaagaatgaagaatgaaga 2448  
 Db 2346 ATTGCAATGAAGAATGGCTGCCATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2405  
 QY 2449 actacotlatacatccagatgagtgactgtgagaaagagcactttacagacaccattgacc 2508  
 Db 2406 ACTACCTATATATCCAGATGGAGTACTGTGAGAGAGCAGCTTTACGAGACACCATGACC 2465  
 QY 2509 agggactgtatcagacacacgctcagactctggaggttttttcgagagatctcggatggat 2568  
 Db 2466 AGGACTGTATCGAGACACCGTCAGACTCTGGAGGCTTTTTCGAGAGATTTCTGGATGGAT 2525  
 QY 2569 tagcttatccatgagaaaggaatgattcaaccgggattttgaagcctgtcaacatttttt 2628  
 Db 2526 TAGCTTATATCCATGAGAAAGGAATGATTACCGGGATTGAAAGCCTGTCAACATTTTTT 2585  
 QY 2629 tggattctgatgacctgtgaaataggtgattttgtgttggcgacagaccatctagcct 2688  
 Db 2586 TGGATTTCTGATGACCATGTGAAATAGGTGATTTTGGTTGGCGACAGACCATCTAGCCT 2645  
 QY 2689 tttctgtgacagcaaacagacgacatcagacgagagacttgattgaagcacccttcag 2748  
 Db 2646 TTTCTGCTGACAGCAAAAGACGATCAGACAGAGACTTGATTAAAGTCAGACCCCTCAG 2705  
 QY 2749 gtcacttaactggatgggttggcaactgctctctctatgtagcccgaggggtcccaaggagca 2808  
 Db 2706 GTCACTTAAGTGGGATGGTTGGCACTGCTCTCTATTAAGCCAGAGGTCCTCAAGGAAGA 2765  
 QY 2809 ccaaatctgtacatacaacagaaagtgtatctctcagctcgtggaattatctcttttga 2868  
 Db 2766 CCAATCTGCATACACCAAGAAAGTGGATCTCTTCAGCCTGGGAATATCTCTTTTGAGA 2825  
 QY 2869 tgtctatcccccattggttcacggcttcagaaaggtatctttgttctcaaccactcagag 2928  
 Db 2826 TGTCTATATACCCCATGGTCACGGCTTCAGAAAGAGTCTTTGTCTCAACCACTCAGAG 2885  
 QY 2929 atccacttcgctaaatttccagaagaacttgcagatgagagacatgcaagcaaat 2988  
 Db 2886 ATCCCACTTCGCTTAAGTTTCCAGAGACTTTGACGATGGAGAGCATGCAAGCAGAAAT 2945  
 QY 2989 cagtcactcctcgtgctgttgaacacgacatccagcaaaacggccacagccacagaaactgc 3048  
 Db 2946 CAGTCATCTCTGCTGTTGAACCCAGATCCACGAAACGGCCACAGCCACAGAGCTGC 3005  
 QY 3049 tcaagagtgcgtgctgccccccacccagatggagagtcagatgcagtgatgaagtgcgc 3108  
 Db 3006 TCAAGAGTGCAGTGTCTGCCCGCCACCCAGATGGAGGAGTCAGAGCTGTCATGAAGTCTGC 3065  
 QY 3109 accacagctgacaaagtcgtgagaaagcctaccgcacacatgagccagatctct 3168  
 Db 3066 ACCACAGCTGCACCAACGTTGGGAAAGCCTTACCACCATGATGGCCCGAGATCTCT 3125  
 QY 3169 cgagcgcatctccctgacctcgattacacctatgacagcgacatgactgaagggcaact 3228  
 Db 3126 CGCAGCGCATCTCCCTGCGCATGATTACACCTATGACAGCGACATCTGAAGGGCACT 3185  
 QY 3229 tctcaatccgtacagcaagatgcagcagatgtgtgtaaacatcactccgcattctta 3288  
 Db 3186 TCTCAATCCGTACAGCAAGATGCAGCAGCATGTGTGTAACCATCATCTCCGATCTTTA 3245  
 QY 3289 aaagacatgagctgttcagttgttactccactactgcttccccgaacacacaaat 3348  
 Db 3246 AAAGACATGAGCTGTTCAGTTGTGTACCTCTGCTTCCCGGAAACAGACAAATAT 3305  
 QY 3349 atgagcacaacagcagctgacctattcatggaacacagcggaatgctgtgtgagtcctcct 3408

Db 3306 ATGAGCACACGAGCTGCCCTATTATCATGACACACAGCGGGATGCTGTGATGCTTCCTT 3365  
QY ttacctcgagatcccttttgcaagatatgtgcaagaataatattattgaattttaaac 3468  
Db 3366 TTGACCTGCGGATCCCTTTTGCAAGATATGTGCAAGAATAATATATTGAATTTAAAC 3425  
QY gactgcatagaaactgtgttgcagccgcaagttagatgatttccccaagaac 3528  
Db 3426 GATACGCATAGAACGTGTGTTCCAGCGCGCAAGTTAGATCGATTTCATCCCAAGAC 3485  
QY ttctgagatgcatattgatattgtcaattctacccaacagctttctgccactgtg 3588  
Db 3486 TTCTGGAGTGTGATTTGATATTGTCACTTCTACCAACAGCTTTCTGCCACTGTG 3545  
QY aattatctacatctatgaataatcatcaagagtttccagacttcagagaaataat 3648  
Db 3546 AATATTATCATATCTATGAATAATCATCAAGAGTTTCCAGCATCTCAGGAAGAAAT 3605  
QY acagtatttttgaaccataccatgtattgaaagcaatactcttcaactgtggatcc 3708  
Db 3606 ACAGTATTTATTGAACCATACCATGTTATTGAAGCAATACTCTTACACTGTGGATCC 3665  
QY cagaagataaactcagtcagtcacattattctgtatgactgtgacagagaagctga 3768  
Db 3666 CAGAAGATAAATCACTCAAGTCTACATTTATCTGTATGATGCTGTGACAGAAAGCTGA 3725  
QY cgaggagagagtggaagctaaattttgtaactgtcttcttctcttaactgtctgtgc 3828  
Db 3726 CGAGGAGAGAGTGAAGCTAAATTTGTAACTGCTTTGCTTCTAATATGCTGTGTC 3785  
QY gactctacaagttattgacagagaaggagatttgcaagatctttatgccacaataaatt 3888  
Db 3786 GACTCTACAAGTTTATTGAACACAGAGGAGATTGCAAGATCTTATGCCAACATAAAT 3845  
QY cattaataaacaagaaaaggttatgcagttgtggaagtgtggttaaaagacctag 3948  
Db 3846 CATTAATAAACAGAAACAGGTTATGCAAGTTGTGGAAGTATGGCTTTAAAGACCTAG 3905  
QY aggaggtgtgtgactgtgaaagaaactcggcaactcaggtctgagctgtgacattgg 4008  
Db 3906 AGGAGGTGTGTGACTGTGGAAGAACTCGGCATCAAGTTACAGGCTTGATCAATTGG 3965  
QY gcttggtttacaaggtgcagcagacacaatggaatcatcttccagttgtggtttcata 4068  
Db 3966 GCTTGGTTTACAAGGTGCAGCAGCACAATGGAATCATCTTCCAGTTGTGGCTTCATCA 4025  
QY acaagagcaaaaggtgtactgaaactcgtcagctggagggcagatatgacctgtga 4128  
Db 4026 AACGAGGCAAAAGGGCTGTACCTGSAATCCTCGACGCTGGAGCAGATATGACCTGTGA 4085  
QY ttcccagtttagagggccacaagctctgggcccagttcccactgcatggggtcagca 4188  
Db 4086 TTCCCAGTTTACAGGGCCACAAAGCTCTGGGCCAGTTCCCACTGCCATTTGGGTCAGCA 4145  
QY tagctatagacaagatatctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 4248  
Db 4146 TAGCTATAGACAAGATATCTGCTGCTGCTCCTCAACATGGAGGAATCTGTTACAATAAGCT 4205  
QY ctgtgacctctgt 4308  
Db 4206 CTGTGACCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4265  
QY cccagaaactctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4368  
Db 4266 CCCAGAACTCTGACAGCAGGCGATCACAGCAAGATCATGTACGACTGGTCACAGTCCC 4325  
QY aagagaattacaagagtagtgcagacataatgaaatcacctatgtggtccctgtgtgt 4428  
Db 4326 AAGAGGAATTTACAGAGTACTGACAGACATCATGAAATCACTATGTGGCCCTGTGTCGG 4385  
QY ataaaagaagaccatgtcaaggttaagttccttctcagagaaagagcagacagagac 4488  
Db 4386 ATAAAGANGAACCCATGTCAAGGTTAAGTCTTTTCAGAGGAAAGGCGACAGAGAGC 4445

QY 4489 gtgtgtgagagactgaactgtgtgaccatgtactgcagaaactgagactaaagtcaactg 4548  
Db 4446 GTGTGCTGGAGACTGAACCTGTGGACCATGTACTTGCAGAACTGAGGACTAAAGTCACTG 4505  
QY atgaagaatgagcagagaagcttccgataatcttgcagtcgcaaaatctgaaggggtcat 4608  
Db 4506 ATGAAGAAATGGCAGAGAAGCTTCCGATATCTTGCAGTGCAGAAATCTGAAGGGGTAT 4565  
QY ttctaatgtctcaggtttgtttgaaatccatgagcaacagtggttcccatgtgagtg 4668  
Db 4566 TTCTTAATGCTTCAGGTTTGTGTTGAAATCCATGAGCAACAGTGGTCCCATTTGTGAGTG 4625  
QY tgcagccccggagaagctgtcagccagcactagagcgctatgaaactcaggtacaaa 4728  
Db 4626 TGCTAGCCCCGGAGAGCTGTACGCCAGCACTAGGAGCGCTATGAAACTCAGGTACAAA 4685  
QY ctgacttcagacacctcccttgcacactatacatcagaagaacagtcgaaattgaaattctgg 4788  
Db 4686 CTCGACTTCAGACCTCCCTTCCCACTTACATCAGAAAGCAGTGAATTTCAATTTCTGG 4745  
QY ctgtgagatcccaaaagaaacaattattacagtttttattcattagatggatgctgatg 4848  
Db 4746 CTGTGATCTACCCAAAGAAACAATATTACAGTTTTTATCATTAGAGTGGATGCTGATG 4805  
QY aacaggcatcttaacacacactgtgaagcagctgtgcacgctgcacgctgcacaaagcaagatacc 4908  
Db 4806 AACAGCATTTAACACAACTGTGAAGCAGCTGTGTGCACGCTGCCCAAGCAAGATACC 4865  
QY tcaattagctgtgatgaatttatacatcaaatgtagaagtagaagaaaggtgtctgtgctat 4968  
Db 4866 TCAATTTAGTCTGTGATGAAATTTTAAACATCAAGTAGAAGAAAGGTGTCTGTGTAT 4925  
QY tctgtacagctatagatgactactacagaatcttattttaacccttaagaaactctcg 5028  
Db 4926 TTCTGTACAGCTATAGAGATGACTACTACAGAACTTATTTTAAACCTTAAAGAACTGTG 4985  
QY ttaacctattcaaacagacagagcttatactggaataatggaatgtgtacattcttc 5088  
Db 4986 TTAACCTCATTTCAACAGACAGAGGCTTATCTGGAATTAATGGAATCTGTACATTATC 5045  
QY ataatttaaaatttaattctcagaagagctgtgtgagtggtcagctcacacctttaatccca 5148  
Db 5046 ATAATTTAAATTTAAATTTCTAAGAGAGGCTGGGTGAGTGGCTCACACCTTTAATCCCA 5105  
QY gactttgggagcgaagcaggaagcagctgtgaaaccagaggttgagaccagct 5206  
Db 5106 GCATTTGGGAAGCCAGGACGAGGAGTCTTGAACACGAGGTTTGAGACCAGCT 5163

RESULT 2  
AB037759  
LOCUS AB037759 4994 bp mRNA PRI 14-MAR-2000  
DEFINITION Homo sapiens mRNA for KIAA1338 protein, partial cds.  
ACCESSION AB037759  
VERSION AB037759.1 GI:7243056  
KEYWORDS  
SOURCE Homo sapiens brain cDNA to mRNA, clone\_lib:pbBluescriptII SK plus clone.fh16948.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (sites)  
AUTHORS Nagase, T., Kikuno, R., Ishikawa, K.I., Hirose, M. and Ohara, O.  
TITLE Prediction of the coding sequences of unidentified human genes.  
XVI. The complete sequences of 150 new cDNA clones from brain which  
code for large proteins in vitro  
JOURNAL DNA Res. 7 (1), 65-73 (2000)  
MEDLINE 20181126  
REFERENCE 2 (bases 1 to 4994)  
AUTHORS Ohara, O., Nagase, T. and Kikuno, R.  
TITLE Direct Submission  
JOURNAL Submitted (31-JAN-2000) to the DDBJ/EMBL/GenBank databases. Osamu







```

Db 3606 TCCCCAGTTTAAAGGGCCACAAAGCTCTGGGGCCAGTTCACCTGCCATGGGTCAGCAT 3665
Qy 4190 agctataagaagaatatctgctgctctccacacagaggaatctgttacaataagctc 4249
Db 3666 AGCTATAGACAAAGATATCTGCTGCTGCTCCACATGAGGAATCTGTTACAATAAGCTC 3725
Qy 4250 ttgtgacctctctggttgaagtggttgctcagatgtctatgtccagggcccatcaacctaac 4309
Db 3726 TTGTGACCTCTGCTGTTGAAGTGTGTCAGATGCTATATGTCCAGGGCCATCAACCTAAC 3785
Qy 4310 ccagaactcttgagacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4369
Db 3786 CCAGAAACTCTGGACACAGGACATCAGCAGAGAAATCATGTACGACTGTGTACAGTCCA 3845
Qy 4370 agaggaattacaagagctactgcagacatcatgataatcattgtacagctgtgcagtcacca 4369
Db 3846 AGAGGAATTACAGAGTACTGACAGACATCATGAAATCATCTATGTCCAGGGCCATCAACCTAAC 3905
Qy 4430 taaagagggaagccatgtcaaggttaagttcttcgagaaggaaaggcagacagagaagcg 4489
Db 3906 TAAAGAAAGGAGGCATGTCAAGGTTAAGTCTTCGAGAGGAAAGGCAGACAGAGAAGCG 3965
Qy 4490 tgtctgagacactgaactgtgacacatgtactgcagaaactgaggaactaaagtcaactga 4549
Db 3966 TGTGCTGGACACTGAACCTTGTGGACCATGTACTGCAGAAACTGAGGACTAAAGTCACTGA 4025
Qy 4550 tgaaggaattgagcagagaagctccgataatcttcagtgcaaaatctgaagggtcatt 4609
Db 4026 TGAAGGAATGCGAGAGAGCTTCGGATATCTTCGAGTGCAGAAATCTGAAGGGGTCAAT 4085
Qy 4610 ttctaattgttcaggttgttgaatccatgtgagcaacagtggtccattgtgagtg 4669
Db 4086 TTTCTATGCTTTCAGGTTTGTGTAATCCATCGAGCAACAGTGGTTCCTCCATTGTGAGTGT 4145
Qy 4670 gctagcccgagaaagctgctcagcagcactagagggcctatgaaactcaggtacaaac 4729
Db 4146 GCTAGCCCGGAGAGCTGCTAGCCAGACACTAGGAGGCGCTATGAACTCAGGTACAAAC 4205
Qy 4730 tcgacttcgaactccctgccaactcatcagaaagcagtgaaattgaaattctggc 4789
Db 4206 TCGACTTCAGACTCCCTTGGCAACTTACATCAGAAAGCAGTGAATTTGAATCTGGC 4265
Qy 4790 tgtggatctaccacaagaacaattacagtttttatcatagagtgagtgatga 4849
Db 4266 TGTGGATCTACCCAAAGAAACAATAATTTACAGTTTTTATCATTAGAGTGGGATGATGA 4325
Qy 4850 acaggcatttaacacaaactgtgaagcagctgctgcacgcctgcacaaagcaagatacct 4909
Db 4326 ACAGGCATTTAACACAACTGTGAAGCAGCTGCTGTACGCTGCCAAGCAAGATACCT 4385
Qy 4910 caaattagctgtgatgaatttaacatcaaatcagaaagcagtgaaattgtaattctgctatt 4969
Db 4386 CAATATTAGTCTGTGATGAATTTTAACTCAATCAAAAGTGAAGAAAGTGTCTGTGCTATT 4445
Qy 4970 tctgtacagctatagagatgactactcagaatcttatttttaacccctaagaagactcgt 5029
Db 4446 TCTGTACAGCTATAGAGATGACTACTACAGAACTCTTATTTTAACCCCTAAAGAACTGCTGT 4505
Qy 5030 taacctattcaaacagacagagcttatactgaaataatggaaattgtgtacattcatca 5089
Db 4506 TAACTCTATTCAACACAGACAGAGCTTATACTGGAATTAATGGAATGTTGTACATTATCA 4565
Qy 5090 taatttaaaataattcttaagaagagctgggtgagtgagtggttcacaccttttaattccag 5149
Db 4566 TAATTTAAATTAATTTCTAAGAAAGAGGCTGGGTGAGTGGCTCACACCTTTAATCCAG 4625
Qy 5150 cactttgggaagcgaagcaggaagactgcttgaacacagaggtttgagacagcctgaag 5209
Db 4626 CACTTTGGGAAGCCAGGAGGAGACTGCTGTAACACGAGGATTTGACACAGCCTGAG 4685
Qy 5210 caacaaagcaagaccatctctataaaaaactaaaaattagttgggcatggtggcaca 5269

```

```

Db 4686 CAACAAAGCAAGACCCCATCTCTATAAAACTAAAAAATTAGTTGGCATGCTGCACA 4745
Qy 5270 tgcctgtagtcctccagctactccagaggtgagatgagatcctcagcctcagaggttga 5329
Db 4746 TGCCTGTAGTCCCGAGCTACTCCAGAGGCTGAGATGATCATCTGAGCCTCAGAGGTTGA 4805
Qy 5330 ggcctgagctgagctgtgactgcgcactgcactccagctcgggacacacagacagacccc 5389
Db 4806 GGCCTGACGTGAGCTGTGACTGCGCCACTGCACCTCAGTCTGGACACACAGACAGACCC 4865
Qy 5390 tgccttaaaaaaagaaaaaataatttttttcttaagaagctgtcctacaaagtga 5449
Db 4866 TGTCTTAAAAAAGAAAAAATAATTTTTTTCTTAAAGAGCTGCTCCTCAACAAAGTTGA 4925
Qy 5450 gctttgttagtttttctcatgtatataataataattatcttttgggataataaatg 5509
Db 4926 GCTTTGTTAGTTTTCATGTGTATATATATATAATTTATCTTTTGGGATATAATAATG 4985
Qy 5510 ctttccatat 5518
Db 4986 CTTTCATAT 4994

```

## RESULT 3

```

HSM802494
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
Source

```

HSM802494 3351 bp mRNA 18-FEB-2000  
Homo sapiens mRNA; cDNA DKFZp434F1312 (from clone DKFZp434F1312); partial cds.  
AL157497  
AL157497.1 GI:7018544  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 3351)  
Duesterhoeft A., Lauber, J., Mewes, H.W., Well, B. and Wiemann, S.  
Direct Submission  
Submitted (18-FEB-2000) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Qiagen (Hilden/Germany) within the CDNA sequencing consortium of the German Genome Project.  
This clone (DKFZp434F1312) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cdna/>.

## FEATURES

```

Source
1..3351
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434F1312"
/clone_lib="434 (synonym: htes3). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
/tissue_type="testis"
1..2819
/gene="DKFZp434F1312"
<1..2819
/gene="DKFZp434F1312"
/notes="similarity to eIF-2alpha kinase (Drosophila melanogaster)"
/codon_start=3
/product="hypothetical protein"
/protein_id="CAB75678.1"
/db_xref="GI:7018545"
/translation="STSGERSASARFPATPGCGSSDDEDDDEHGGVFSQSFSPASD
ESDIIPDNEDENSKQNDQCNKNGCHESEPSVTTEAVHYLIQMEICEKSTLRDT
IDQGLYRDTVRLWRLRFREILDSGLAYIHEKGMIRHOLKPNVIFLSDSDHRTGLAT
HLAFSADSKDDQDTGLKSDPSGHLTGMVGTALYSPVQSGTSAKYNQKVDLFLS
GLIFFEMSYHPMTASERIFVLNQLRDTSPKFPEDFDGGEHAKQKSVISWLLNHDA

```

KRPTATELLKSELLPPQOMESEHELVUHTLTNNVDGKAYTMMQAQIFSORISPAIDI  
 YTSYDILKGNFSRDTAKMOQHCYETIIRIFKXGVAOLCTPLLRQRIQYEHNEAAL  
 FMDJSGNMLPDIAPRYAVARNINLNKRYEIERVPRKDLRPHKLECCAF  
 DIVTNSGTFPABEITVYIEIOBFPALQERNYSVINHTMLKAILLCGCPEDK  
 LSQVYIILADVATEKLTREVEKACNLSNSNLSCLRYKEIQOGLQDMITPDL  
 IKOKTIGIAQLVKYKDLDEEVLGKLGKLOVILNGLVKYQOQHIGYIFOFVART  
 KRQRVPEILLAGSGMYRDLILPQFRQNALGPVPTAGVSTAIDKAAAVLNMEESTV  
 ISSCDILVSVGOMSGRDLITNQLKLTAGTAEITMDWSQSOSELOEYGRCHIEITV  
 ALKDSGKSHVYKFEKEQRQETKRVLETFLVDHVLQKLTARTVDYDRNGREASDNTAV  
 QNLKSGFVSNAGLVEFTHGATVPYVSUVAPELKSZTRRRYETQVOTQLOQTISLANLHQ  
 KSSBETIEALVDLPEKETIQGLFSLNDEDAEQEFTTVTKOLLSRLRQRYLKLVCDEITYN  
 IEKVEKYSVLFLEYSYRDDYRILF"

polvA signal

polya\_site  
3331

POLYA\_SIZE 1017 a  
BASE COUNT

BASE C  
ORIGIN

Query Match 60.3%; Score 3332; DB 93; Length 3351;  
Best Local Similarity 100.0%; Pred. NO. 0;  
Matches 3332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	2194	ggagcacttcggcgagcgctcggcagctgcgcgtttccocggccacccggcccggggtcca	2253
Db		1 GGAGCACTTCGGGCGAGCGCTCGGCAGTGCOCGCTTTCCCGCACGGCCGGGGTCCA	60
QY	2254	gcatgacgaggagcagcaacagagcagcagctggcgtctctccagctcctcctgc	2313
Db		61 GCATGACGAGGACGACGAGGAGGACGACGGTGGCGTCTTCTCCAGTCTCTCTGC	120
QY	2314	ctgttcacgattcgaaagtattattctttggacaatgagaatgagacagtaaaagtc	2373
Db		121 CTGCTTCAGATCTCGAAATGATATTATCTTTGACAATGAAGATGAGAACAGATAAAGTC	180
QY	2374	agaaatcagatgaagattccaatgaaaagaatggctgccatgaaatgagccatcagtga	2433
Db		181 AGAATCAGGATGAAGATTCGATGAAGAAAGATGGCTGCCATGAAGATGAGCCATCAGTGA	240
QY	2434	cgactgagctgtgcactacctatcatccagatggagtctgagaaagagcactttac	2493
Db		241 CGACTGAGGCTGTGCATCTACCTATACATCCATCGATGGAGTACTGTGAGAAGACATTTTAC	300
QY	2494	gagacacattgaccaggagctatccagacacacgctcagactctggagggcttttcgag	2553
Db		301 GAGACACCATTTGACCAGGGGACTGTATCGAGACACCGTTCAGACTCTGGAGGCTTTTTCGAG	360
QY	2554	agattctggatggattagctttatccatgagaaaggaaatgattccocggagatttgaagc	2613
Db		361 AGATTCTGGATGGATTAGCTTTATTCATGAGAAAGGAATGATTCACTGGGATTTGAAGC	420
QY	2614	ctgtcaacatttttttgattctgatgaccatgtaaaatagggtgattttggttggcga	2673
Db		421 CTGTCAACATTTTTTTGGATCTGTATGACCATCTGAAAATAGTGTGATTTGGTTTGGCGGA	480
QY	2674	cagaccaactagccttttctgtgcagcaacaacagcagatcagacagagagacttgatta	2733
Db		481 CAGACCATCTAGCCTTTTCTGCTGACAGCAAAACAGACGATCAGACAGAGACTTGATTA	540
QY	2734	agtcagacccttcaggctcaacttaactggagatgtggcactctctctatgtaaagccag	2793
Db		541 AGTCAGACCCCTTCAGGTCACCTTAACCTGGGATGTTGGCACTGCTCTCTATGTAGCCCGAG	600
QY	2794	aggttccaagggaagcaccacaaatcttgatcaccaaccagaaatggatctcttcagcctggaaa	2853
Db		601 AGGTCCAAGGAAGCACCAAAATCTGCATACAAACAGAAAGTGGATCTCTTCAGCCCTGGGAA	660
QY	2854	ttatctcttttgaagatgctcctatcaacccatgttcacgacttcagaaagactcttcttc	2913
Db		661 TTATCTCTCTTTGAGATGTCCTATCACCCCATGGTCAAGGCTTCAGAAAGCATCTTTGTTC	720
QY	2914	tcaaccaactcagatcccacttcgctaaagttccagaagaactttgacatggagagc	2973
Db		721 TCATCCAACTTCAGATCTCCACTTCGCTTAAGTTTCAGAAAGACTTTGACGATGGAGAGC	780



QY 4054 ttgtgctttcatcaaacgaaggcaagcgtgtacotgaaatcctcgcagctggaggca 4113  
|||||  
Db 1861 TTGTGGCTTTTCATCAAGCAGGCAAGGGCTGTACCTGAAATCCCTCGCAGCTGGAGGCA 1920  
|||||  
QY 4114 gatatacctctgattccccagtttagagggccacaagctctggggccaggtcccaactg 4173  
|||||  
Db 1921 GATATGACCTGCTGATTCCCGCAGTTTAGAGGGCCACAAGCTCTGGGCCAGTTCACACG 1980  
|||||  
QY 4174 ccattggggtcagcatagctatagacaagaatctctgctgctcctcaacatggaggaat 4233  
|||||  
Db 1981 CCATTGGGTCAGCATASCATATAGCAAGATATCTGCTGCTGCTCTCAACATGGAGGAAT 2040  
|||||  
QY 4234 ctgtacaataagctctgtgacctcctgtgttaagttgttgctgagatgctctatgtcca 4293  
|||||  
Db 2041 CTGTTACAATAAGCTCTGTGACCTCCTGTTGTAAGTGTGGTCAGATGCTATGTCCA 2100  
|||||  
QY 4294 gggccatcaacctaaaccagaaactctggacagcaggcctcacagcaaaatcatgtacg 4353  
|||||  
Db 2101 GGGCCATCAACCTAACCCAGAACTCTGGACAGCAGGCAATCACAGCAAAATCATGTAGC 2160  
|||||  
QY 4354 actgggtcacagtcctcccaagggaattacaagagctactgcagacatcatgaaatcacctatg 4413  
|||||  
Db 2161 ACTGGTCACAGTCCCAAGAGGAATTACAGAGTACTGTCAGACATCATGAAATCACCTATG 2220  
|||||  
QY 4414 tggccctgtctcggaataaagaaggaagccatgtcaaggttaagctcttcgagaaagaaa 4473  
|||||  
Db 2221 TGGCCCTGTCTCGGATAAAGAAAGGAGCCATGTCAAGGTTAAGTCTTTCGAGAAGGAAA 2280  
|||||  
QY 4474 ggcagacagagagcgtgtctggagactgaactgtgacacatctactgcagaaactga 4533  
|||||  
Db 2281 GGCAGACAGAGAAGCTGTGCTGGAGCTGAATCTGTGGACCATGTACTGCAGAAACTGA 2340  
|||||  
QY 4534 ggaactaaactactgatgaagaagaaatggcagagaagcttcgcataatctgtcagtgcaaa 4593  
|||||  
Db 2341 GGCTAACTACTGATGAAGAGGAATGGCAGAGAAGCTTCGATAAATCTTGCACTGCAAA 2400  
|||||  
QY 4594 atctgaagggttcatttcttaagtcttcagctgtgtgttgtaaatccatggagcaacagtgg 4653  
|||||  
Db 2401 ATCTGAAGGGGTCAATTTCTAATGCTTCAGGTTTGTGTGTAATCCATGGACCAACAGTGG 2460  
|||||  
QY 4654 ttccattgtgaagtgtgtagcccggaagcgtgtcagcagcagcactgagggcgctatg 4713  
|||||  
Db 2461 TTCCCATTTGTGAGTGTGTAGCCCGGAGAAAGCTGTGAGCAGCAGCTAGGAGCGCTATG 2520  
|||||  
QY 4714 aaactcaggtacaaactcagctcagacctccttgccaacttacatcagaaaagcagtg 4773  
|||||  
Db 2521 AAACTCAGGTACAACTGCACTTCAGACCTCCCTTGCCAACTTACATCAGAAAAGCAGTG 2580  
|||||  
QY 4774 aaattgaattcgtgtggtatctacccaaagaaacaaatattacagtttttatcattag 4833  
|||||  
Db 2581 AAATTGAAATTTCTGGCTGTGGATCTACCCAAAGAAACAATATTACAGTTTTATCATTTAG 2640  
|||||  
QY 4834 agtgggactgctgaagcagcatttaacacactgtaagcagcgtgtgtcagcgcctgc 4893  
|||||  
Db 2641 AGTGGGATGCTGTGAACAGGCAATTTAACACACTGTGAAGCAGCTGTGTCAGGCGCTGC 2700  
|||||  
QY 4894 caaagcaagatacctcaaatagctgtgatgaatttataacatcaagtagaaaaaa 4953  
|||||  
Db 2701 CAAAGCAAGATACCTCAATATTAGTCTGTGATGAAATTTATACATCAAAAGTAGAAAAA 2760  
|||||  
QY 4954 aggtgtctgctattctctgacagctatagagatgactactacagaattcttttttaac 5013  
|||||  
Db 2761 AGGTGTCTGTGCTATTCTGTACAGTATAGAGATGACTACTACAGAAATCTTATTTTAAC 2820  
|||||  
QY 5014 cctaaagaactcgtttaacctcatcaacagacagagcgttactggaataatgaa 5073  
|||||  
Db 2821 CCTAAAGAACTGTCGTATACCTCATTAACACAGACAGAGGCTTATACCTGGAATATGAA 2880  
|||||  
QY 5074 tgtgtacatcatcaataatttaaaatttaattcttaagaagggctgggtcagtggtc 5133  
|||||  
Db 2881 TGTGTGACATTCATCAATAATTTAAATTTAAATTTCTAAGAGAGGCTGGGTGCTGCTGCTC 2940  
|||||  
QY 5134 acacctttaatcccgacactttgggaagccaaaggcaggaagactgcttgaaccaggagt 5193  
|||||

Db 2941 ACACCTTTAATCCAGCACCTTTGGGAAGCCAGGCAAGAGAGCTGCTTGAACACGAGGT 3000  
|||||  
QY 5194 ttgagcagcctgagcaacaaagcaagaccccatctctataaaactaaaaaattagt 5253  
|||||  
Db 3001 TTGAGACAGCCTGAGCAACAAAGCAAGACCCCATCTCTATAAAACTAAAAAATAGT 3060  
|||||  
QY 5254 tgggcattggtgacacatgctgttagtccacagctactccagagctgagatgcatctg 5313  
|||||  
Db 3061 TGGCATGTTGGCACATGCTGTAGTCCAGAGCTACTCCAGAGCTGAGATGATCATCTG 3120  
|||||  
QY 5314 agcctcagggaggttgaggctgagctgagctgtgactgagcagcactcagctcagctgga 5373  
|||||  
Db 3121 AGCCTCAGGAGGTTGAGGCTGAGCTGAGCTGTGACTGCCCACTGCACCTCCAGTCTGGGA 3180  
|||||  
QY 5374 caacagcagcagccctgctcttaaaaaaagaaaaaatttttttctaagaagc 5433  
|||||  
Db 3181 CAACAGCAGCAAGACCCCTGCTTTAAAAAAGAAAAAATTTTCTTCTAAGAAGC 3240  
|||||  
QY 5434 tgtctacaaaagtgtgagctttgttagtttttcatgtgtaatatataaaatttatcttt 5493  
|||||  
Db 3241 TGTCTTACAAAGTTGAGCTTGTAGTTTTCATGTGTATATATATAAATTTATCTTT 3300  
|||||  
QY 5494 tgggataataaagtcttctcatatcacctgca 5525  
|||||  
Db 3301 TGGGATATAATAAATGCTTTTCATATACCTGCA 3332  
|||||

## RESULT 4

AK027011 LOCUS AK027011 2986 bp mRNA PRI  
DEFINITION Homo sapiens CDNA: FLJ23358 fis, clone HEP14996. 29-SEP-2000  
ACCESSION AK027011  
VERSION AK027011.1 GI:10440017  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens hepatoma cell\_line:HEPG2 CDNA to mRNA, clone\_lib:HEP  
clone:HEP14996.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (sites)  
Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,  
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,  
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.  
NEDO human cDNA sequencing project  
Unpublished (2000)  
2 (bases 1 to 2986)  
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,  
Shibahara,T., Tanaka,T. and Nakamura,Y.  
Direct Submission  
Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio  
Sugano, Institute of Medical Science, University of Tokyo,  
Laboratory of Genome Structure Analysis, Human Genome Center;  
Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan  
(E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286,  
Fax:81-3-5449-5416)  
NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan; cDNA full insert  
sequencing: Research Association for Biotechnology; cDNA library  
construction, 5'- & 3'-end one pass sequencing; Department of  
Virology and Human Genome Center, Institute of Medical Science,  
University of Tokyo (partly supported by Science and Technology  
Agency).  
FEATURES  
Location/Qualifiers  
1..2986  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="HEPG2"  
/cell\_type="hepatoma"  
/clone="HEP14996"  
/clone\_lib="HEP"  
/note="cloning vector pME18SFL3"  
40..2460  
CDS

/note="unnamed protein product"									
/codon_start=1									
/protein_id="BAB15625.1"									
/db_xref="gi:10440018"									
/translation="MIHRDLKPVNIPLDFDDHKVIGDFGLATDHLAFSADSKODDGTG ERIFVLNDRPSTSPKEPFEDFGEHAKOKSVISLWNHDPKAPRTATELLKSELLP DLIKSDPSGLTGMVGLKALVSPDEGOKTSAYNOKVDFLSGLGIFFEMSYHPMVTAS POMESELHEVHHTLNVGDKAYPMMAQIFSORISPAIDYTDSDILKGNESIRTA KMOQHVCTIIIRIFKRIGAVTCLPPLLRLNRQIYEHNEALFMDHSGLMLVLPFLRAI IPFAYHVNILNLKPKRYCIVELFRDRLPHRKLLECEAFDITVSTTNSPFLPAEI ITYTVEIQEPFALQERNYSIYNLHMTLLKAILLHCGIPEDKLSQYIILLYLDVATEKL TRREYAEFCYLSNSLGRILYFTQERQDQDLMTPLINSLIKQRTGIAOLKVLGKLK DLEEVGGLLKGLKLOVILNLGLVATKVQHQHNGIIFQFVATKRRQAVPEILAAGR YDLLIPQRGLKGLVPTAIGVSIADKISAAVLNMBESVTISSCDLLVYVSGQMSM SRAINLTKLMTAGTGAIEIMDMSQSQEOLVGRHRHEITVVALVSKGFSHVKVSFI EKERQTVPIETELVHVLQKLTAVTDENRGREASDNLAVNLKSGFSNASGLPEI HGATVPIVSVLAPKLSASRRRYENFQVTRLOPSTLANLHOKKSETEILLAVDLPKET ILQFSLSEWDNDEQAFNTVKNQLSLRUPKORYLKLVCDEIYNIKVERKVSVLFLYSVR DDYIRILF"									
BASE COUNT	921 a	641 c	663 g	761 t					
ORIGIN									
Query Match	48.0%;	Score 2650;	DB 89;	Length 2986;					
Best Local Similarity	99.9%;	Pred. No. 0;							
Matches 2850;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;	
y	2553	qaagattcggatgattagcttatatccatgagaagaatgattaccacgggatttgaag	2612						
b	1	GAGATTCGGATGAGTTAGCTTATATCCATGAGAAAGAAATGATTCACCGGATTTGAAG	60						
y	2613	ccgtgcaacatttttggattctgtgaccatgtgaaaaatgattgatttggcttggcg	2672						
b	61	CCTGTCAACATTTTTTGGATTTGTATGATCACCATGTGAAAAATAGGTGATTTTGGTTGGCG	120						
y	2673	acagaccatctagcctttctgtgcagcagcaacaagaacgacgcagacgagagacttgatt	2732						
b	121	ACAGACCATCTAGCCTTTTCTGCTGCAGCAACAAAGACGATCAGACAGGAGACTTGATT	180						
y	2733	aagtcagacccttcaggctaacctggatggatggttgcaactgctctctatgtaagccca	2792						
b	181	AAGTCAAGACCCCTTCAGTGCACCTTAACCTGGATGGTTGGCACCTGCTCTATGTAAAGCCCA	240						
y	2793	gaggtccaaagggaagcaccaaatctgtacataccacagaagaagtggatctcttcagcctggga	2852						
b	241	GAGGTCCAGGAAGCAACCAAACTGCATACAACACAGAAAGTGGATCTCTTCAGCGCTGGGA	300						
y	2853	attatcttcttgagatgctcatcaccatggtcacggtctcagaagatcttggcttggct	2912						
b	301	ATTATCTCTTTGAGATGTCCTATACCCCATGGTACGGCTTCAGAAAGGATCTTTGTT	360						
y	2913	ctcaaccaactcagagatcccaacttcgctaaagttccagaagaacttgaogatggagag	2972						
b	361	CTCAACCAACTCAGAGATCCCACTTCGGCTTAAGTTCAGAGACATTTTCACGATGGAGAG	420						
y	2973	catgcaagcagaaaaatcagtcacatctcgtgttgaaacacgacatccagcaaaacgccc	3032						
b	421	CATGCAAGCAGAAAATCAGTCAATCTCTGGCTGTTGAACCAACGATCCACAAACGCGCC	480						
y	3033	acagccacagaactgctcaagagtgcgtgcctcccccaccccgatgagagatcagag	3092						
b	481	ACAGCCACAGAACTGCTCAAGAGTGAAGTGTGCTGCCCCACCCAGATGGAGAGTCAAG	540						
y	3093	ctgcatgaagtgtgcacacacgctgaccaactgcatggaggggaagcctaccgcgaccatg	3152						
b	541	CTGATGAAGTGTGCACACACGCTGCACCAACCTGGAATGGGAAGGCTTACCGCACCATG	600						
y	3153	atgggccagatcttcttcgcagcgcacatctccctgccatcgattaccacatgaacgcgac	3212						
b	601	ATGGCCCGAGATCTCTCGCAGCGCATCTCCCTGCGCATCGATTACACCTATGACAGCGAC	660						
y	3213	atactgaagggaactctcaatccgtacagcagaagtgcagcagcatgtgtgtgaacc	3272						

QY 4353 gactgtcacagtcccaagagaataacaaagtagtactcagacatcatgaaatcacctat 4412  
 DB 1801 GACTGGTACAGTCCCAAGAGAAATTAACAAGATGACTCAGACATCATGAATCACCTAT 1860  
 QY 4413 gtgcccctgtctcgataaagaagaagccatgtcaagggttaagtcttcgagaagaa 4472  
 DB 1861 GTGGCCCTGTCTCGGATAAAGAAGGAGCCATGTCAAGGTTAAGTCTTTCGAGAAGAA 1920  
 QY 4473 aggcacagagaagcgtgtgctggagactgaactgtgagccatgtactcagaaaactg 4532  
 DB 1921 AGGCAGACAGAGAAGCGGTGTCTGGAGACTGAACCTGTGGACCATGTACTCGAAAACTG 1980  
 QY 4533 agdactaaagtcaactgatgaagaagatggcagagaagcttcccgataatcttgcagtcac 4592  
 DB 1981 AGGACTAAGTCACTGATGAAGGATGGCAGAGAAGCTTCCGATATATCTTGCAGTGCAA 2040  
 QY 4593 aatctgaagggttcattcttaagtctcaggtttgttgaataatccatggagcaacagt 4652  
 DB 2041 AATCTGAAGGGGTCAATTTCTAATGCTTCAGGTTGTGTAATCCATCGGAGCAACAGTG 2100  
 QY 4653 gtcccatgtgagtgctagcccgagagagctgtcagccagcactagaggcgctat 4712  
 DB 2101 GTTCCATGTGAGTGTGCTAGCCCGGAGAGCTGTGAGCAGCAGTGTAGGAGCGCTAT 2160  
 QY 4713 gaaactcaggtacaaactcgacttcagacctcccttgcacacttcatcatcagaaaagcagt 4772  
 DB 2161 GAAACTCAGGTACAACTCGACTTCAGACCTCCCTGCCACCTTACATCATGAAAGCAGT 2220  
 QY 4773 gaaattgaaattctggctgtggtatcccaagaagaacaatatattacagttttatcata 4832  
 DB 2221 GAAATTGAAATTTCTGGCTGTGGATCTACCCAAAGAAACAATATTACAGTTTTATCATTA 2280  
 QY 4833 gagtggagatgctgatgaacaggtattacaacactgtgaagcagctgtcagcgcctg 4892  
 DB 2281 GAGTGGATGCTGATGACAGGCATTTAACAACACTGTGAAGCAGCTGTCTCACGCGCTG 2340  
 QY 4893 ccaagcaagaatcacctcctaattagctgtgtgatgaatttataacatcaaatgagaaaaa 4952  
 DB 2341 CCAAGCAAGATACCTCAAAATAGTCTGTGATGAATTTAATACATCAAGTAGAAAAA 2400  
 QY 4953 aagtgctgtgctattctgtacagctatagagatgactactacagaactctatttaa 5012  
 DB 2401 AAGTGTCTGTGCTATTCTGTACAGCTATAGAGATGACTACTACAGAATCTTATTTTAA 2460  
 QY 5013 ccctaaagaactgcgttaacctattcaacagacagagagcttatactaggaataatgga 5072  
 DB 2461 CCCTGAAGAACTGTCATTAACTTCAACACAGACAGAGGCTTATCTGGAATATATGGA 2520  
 QY 5073 atgtgtacattcatataatttaaaatttaaaattcttaagaagaggtgggtcagtgagct 5132  
 DB 2521 ATGTTGTACATTTCATATAATTTAAATTTAAATTTCTAAGAAAGAGGCTGGGTGCGTGGCT 2580  
 QY 5133 cacaccttaatccagcactttggaagccaaagcaggaagactgcttgaaccaggag 5192  
 DB 2581 CACACCTTTAATCCACACTTTGGGAAGCCAGGAGGAAAGACTGCTGAAACCAAGGAG 2640  
 QY 5193 ttgagaccagcctgagcaacaagaagaccctctctataaaactataaaataatag 5252  
 DB 2641 TTTGAGACCAGCCTGAGCAACAAGCAAGACACCCCATCTCTATTAATAACTTAAAAAATAG 2700  
 QY 5253 ttgggaatgtggcacatgctgtagtcacagctactccagaggtcagatgagatcatct 5312  
 DB 2701 TTGGGCATGTGGACATGCTGTAGTCCAGCTACTCCAGAGGCTGAGATGATCATCT 2760  
 QY 5313 gagcctcagaggtgtgaggtcagtgagctgtgactgcagcactgcactccagctcgtgg 5372  
 DB 2761 GAGCCTCAGAGGTTGAGGCTGAGTGTGACTGTGGCCACTGCATCCAGTCTGGG 2820  
 QY 5373 acaagagcaagaccctgtcttataaaaaaa 5406  
 DB 2821 ACAAGAGCAAGACCTGTCTTAAAAAATAA 2854

## RESULT 5

HSA243428 2162 bp mRNA PRI 15-OCT-1999  
 LOCUS Homo sapiens partial mRNA for putative eIF2 alpha kinase (GCN2 gene)  
 DEFINITION  
 ACCESSION AJ243428  
 VERSION AJ243428.1 GI:6065913  
 KEYWORDS eIF2 alpha kinase; GCN2 gene.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 2162)  
 AUTHORS Berlanga, J.J., Santoyo, J., and De Haro, C.  
 TITLE Characterization of a mammalian homolog of the GCN2 eukaryotic initiation factor 2alpha kinase  
 JOURNAL Eur. J. Biochem. 265 (2), 754-762 (1999)  
 MEDLINE 99435990  
 REFERENCE 2 (bases 1 to 2162)  
 AUTHORS Santoyo, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-JUN-1999) Santoyo J., Molecular Biology, Centro de Biologia Molecular 'Severo Ochoa', Fac. Ciencias. U.A.M., Cantoblanco, Madrid, 28049, SPAIN  
 FEATURES  
 Location/Qualifiers  
 1..2162  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 1..2144  
 /gene="GCN2"  
 <1..1649  
 /gene="GCN2"  
 /function="protein synthesis control"  
 /codon\_start=3  
 /product="putative eIF2 alpha kinase"  
 /protein\_id="CA58360.1"  
 /db\_xref="GI:6065914"  
 /translation="ALFMDHSGMLVLPFDLRIFARVAVARNILILKRYCIERVFRP  
 RKLDRFPKLELCEADIVTSTSTFLPTAEIITYIEIIEFFALQERNYSIYLNHT  
 MLKAILLHCIGIPEDKLSQVYIILYDAVTEKLRREVAKFCNLSLSNSLRLYKFI  
 BOKGLQDLMTPTINSIKQVTKGLVGLKLEEVVGLKLGILKLOVLINLGLVY  
 KVOHNGIIFQVAFIKRRRAVPEILAAAGGRYDLIPQFRPOLGVPPTAIGVSTA  
 IDKTSAAVLNMEESTVITSSCDLLVSVGMSRAINLTQKLTAGTAGTAELMWDSQS  
 QEELEQYCRHHEIITYVALVSKESHVKVSKFEKERQTKRVLETELVDHVLQKLRK  
 VIDENKREASONLAVQNLKGSFNASGLFEHGATVVIIVSLAPEKLSARRRYE  
 TOVOTRIQTSANLHQKSSSEIILAVDLPKETILQFLSLEWDADAQAFNTVTKOLLSR  
 LPKORYLKLVCDEIYINIKVERKVSFLYSYRDDYVYRLF"  
 <2139..2144  
 /gene="GCN2"  
 polyA\_signal  
 BASE COUNT 672 a 434 c 481 g 575 t  
 ORIGIN  
 Query Match 38.2%; Score 2111; DB 93; Length 2162;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2161; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3364 ctgcacctatctgacacagcgagctgtgtgatgctctcttgcacctgcgagatcc 3423  
 DB 1 CTGCCCTATTTCATGACCACAGCGGATGCTGCTCTCTTTTGACCTCGGATCC 60  
 QY 3424 ctttgcagaatagtgcgaagaaataatatattgaatttaaacagactctgagtgatgc 3483  
 DB 61 CTTTTCGAAGATATGTGGCAGAAATAATATTATTATTTTAAACGATACGTAGAAC 120  
 QY 3484 gtgtgtcagcgccgcgaagttagatcgatttcacccaaagaaactctctgagtgatgc 3543  
 DB 121 GTGTGTTTCAGGCGCGCAAGTTAGATCGATTTCATCCCAAGAACTTCTGGAGTGTGCAT 180  
 QY 3544 ttgatattgtcactctctaccaccacagcttctgcccactgctgaaattatctacacta 3603  
 DB 181 TTGATATTGTCACTTCTACCACCAACAGCTTCTGCCACTGCTGAAATTTATCTACACTA 240

QY 3604 tctatgaatcatcccaagagcttccagcaacttcagaaagaaattacagattattttga 3663  
Db TCTATGAATCATCCCAAGAGCTTTCCAGCACTTCAGGAAAGAAATACAGTATTTATTTGA 300  
QY 3664 accataccatgctatgaagcaactctctacactgtgggacccagagaataaactca 3723  
Db TCTATGAATCATCCCAAGAGCTTTCCAGCACTTCAGGAAAGAAATACAGTATTTATTTGA 300  
QY 301 ACCATACCATGTTATTTGAAGCAATCTCTTACACTGTGGGATCCAGAGATAAATCTCA 360  
Db TCTATGAATCATCCCAAGAGCTTTCCAGCACTTCAGGAAAGAAATACAGTATTTATTTGA 300  
QY 3724 gtcaagtctacattattctgtatgctgtgacagagaagctgacagagagaagtg 3783  
Db GTCAAGTCTACATTTATTTCTGTATGCTGTGACAGAGAAAGCTGACAGGAGAGAAAGTGG 420  
QY 3784 aagtaaatatttgaactgtcttcttcttaataagctgtgtcgaactctcaagattta 3843  
Db AAGCTAAATTTGTAATCTGTCTTTGCTTCTAATAGTCTGTGCTGACTCTCAAGATTTA 480  
QY 3844 ttgaacagaagggagatttgaagatcttgcacaaataaaatttcaataaaacaga 3903  
Db TTGAACAGAGGGAGATTTCAGAGATCTTATGCCARCAATAAATTCATTAAATAAACAGA 540  
QY 3904 aaacaggtattgcacagttggtgaagtatggtcttaaaagacctagagaggttcttggac 3963  
Db AAACAGGTTATTCACAGTGTGTAAGTATGGCTTAAAGACCTAGAGGAGGTTGTTGGAC 600  
QY 3964 tgttgaagaactcggcatcaagttacagctcttgatcaatttggccttggtttacaag 4023  
Db TGTGAAGAAACTCGGCATCAAGTTACAGGCTTGTGATCAATTTGGGCTTGGTTTACAGG 660  
QY 4024 tgcagcagcaaatgggaatcatctccagtttgggttttcaaaacgaagcaaggg 4083  
Db TGCAGCAGCACAAATGGAATCATCTCCAGTTGTGGCTTTCATCAACAGCAAGCAAGGG 720  
QY 4084 ctgtacctgaatcctcgcagctggagcagatgatgacctgtgattccccagtttagag 4143  
Db CTGTACCTGAAATCCTCGACGTGGAGGAGATATGACCTGCTGATTCCCGCAGTTTAGAG 780  
QY 4144 ggcacaagctctggccagcttccactgcatggtgggtcagcatagctatagacaaga 4203  
Db GGCCACAAGCTCTGGGGCCAGTTCCTACTGCCATTTGGGGTCAGCATATAGTATAGACAAGA 840  
QY 4204 tatctgctgctgtcctcaactgaggaatctgttacaataagctctgtgacctctgg 4263  
Db TATCTGCTGCTCCTCAACATGAGGAAATCTGTTACAATAAGCTCTGTGACCTCTGCG 900  
QY 4264 ttgtaagttgttgcagatgctatgtccagggccatacaaccttaaccagaaactctgga 4323  
Db TTGTAAGTGTGTTGGTCAGATGCTATGTCCAGGGCCATCAACCTAACCCAGAACTCTGGA 960  
QY 4324 cagcagcatcacagcagaataatcatgtacactgtgtcacagtcaccaagaggaattacaag 4383  
Db CAGCAGGATCATCAGCAGAAATCATGTACGACTGGTCACAGTCCCAAGAGGAATTTACAAG 1020  
QY 4384 agtactgcagacatcatgaaatcacctatgtggcccttgtctcggataaagaagaagcc 4443  
Db AGTACTGCAGACATCATGAAATCACCTATGTGGCCCTTGTCTCGGATAAAGAGGAAGCC 1080  
QY 4444 atgtcaaggttaagcttctcagaagaagaaagcagacagaagcgtgctgagagactg 4503  
Db ATGTCAAGGTTAAGTCTTTTCGAGAGAAAGGACAGAGAGAGGCTGTGCTGGAGACTG 1140  
QY 4504 aacttgtgaccatgactgcagaaactgagactaaagtacactgatgaaaggaattggca 4563  
Db AACTTGTGGACCATGACTGCAGAAACTGAGGACTTAAAGTCACTGATGAAAGGAATGGCA 1200  
QY 4564 gagaagctccogataatcttgcagtgcaaaatctgaaaggggtcattttcattatcttcag 4623  
Db GAGAAGCTTCCGATATCTTGCAGTGCAAAATCTGAAGGGGTCAATTTCTAATGCTTCAG 1260  
QY 4624 gtttcttgaatccatgcagcaacagtggttcccatgtgagtgctgtagccccggaga 4683  
Db GTTTGTTGAAATCCATGCAGCAACAGTGGTTCCTCATTTGTGAGTGTGTAGCCCGGAGA 1320  
1261 GTTTGTTGAAATCCATGCAGCAACAGTGGTTCCTCATTTGTGAGTGTGTAGCCCGGAGA 1320

QY 4684 agctgtcagcagcactagaggcgctatgaaactcagggtacaaactcagacttcagacct 4743  
Db ACTGTACGACGACACTAGGAGCGCTATGAACTCAGGTACAACTCGACTTCAGACTT 1380  
QY 4744 cccctgcacacttacatcagaaagcagtgaaattgaaattcctgctggtgagctaccca 4803  
Db CCCCCTGCCAACTTACATCAGAAAGCAGTGAATTTGAATTTCTGGCTGTGGATCTACCCA 1440  
QY 4804 aagaacaataattacagtttttattcattagagtggtgagtgctgagaaagcagcatttaaca 4863  
Db AAGAAACAATATTACAGTTTTTATCATTTAGAGTGGATGCTGATGAACAGCAATTTAACA 1500  
QY 4864 caactgtgaagcagctgtctcaogctgcaccccaagcaagatacctcaaatagctctgtg 4923  
Db CAACCTGTGAAGCAGCTGTCTCACGCCCTGCCAAAGCAAGATACCTCAAAATAGCTGTG 1560  
QY 4924 atgaataattatacaatacaagtagaaaaaaggtgtctgtctatttctgtacagctata 4983  
Db ATGAATTTTATTAACATCAAAAGTAGAAAAAAGGTGCTGTGCTATTCTGTACAGCTATA 1620  
QY 4984 gagatgactactacagaatctttatttaacccctaaagaaactgtcgttaacctcattcaaa 5043  
Db GAGATGACTACTACAGATCTTATTTTAAACCTTAAAGAACTGTCGTTAACCTCATTTCAA 1680  
QY 5044 cagacagaggtcttactggaaataatggaaatgtgtacattoatcatataatttaaaattaa 5103  
Db CAGACAGAGGCTTATACTCGAATAATGGAATTTGTACATTTATTTAAATTTAA 1740  
QY 5104 attctaaagaaggtgtgtgagtggtgagtggtgagtggtgagtggtgagtggtgagtggt 5163  
Db ATTCTAAGAAGAGGCTGTGGTGTGAGTGGCTTACACCTTTTAAATCCCAAGCACTTTGGGAAGCC 1800  
QY 5164 aaggcaggaagactgcttgaacaccagaggtttgagaccagcctgagcaacaaagcaagac 5223  
Db AAGCAGGAGGAGACTGCTTTGAAACCCAGGAGTTTGAGACCAGCTGAGCAACAAAGCAAGAC 1860  
QY 5224 cccatctctataaaactaaaaaattagttgggcatggtgggcatggtggcagctgagtgagct 5283  
Db CCCATCTCTATAAAAACTAAAAAATTTAGTTGGCATGCTGGCAGCATGCTGTAGTCCCA 1920  
QY 5284 gctactcagagctgagatgagctgagctcagagaggttgaggtcagtgagct 5343  
Db GCTACTCCAGAGCTGAGATGATGATCTGAGCCTCAGGAGGTTGAGGCTGAGTGTAGCT 1980  
QY 5344 gtagctgcgcccactgcactccagctgggacacacagagcaagacccctgctcttaaaaaaa 5403  
Db GTGACTGCGCCACTGCACTCCAGTCTGGGACAAACAGAGCAAGACCCCTGTCTTTAAAAAAA 2040  
QY 5404 aaagaaaaaatt 5463  
Db AAAGAAAAAATTTTCTTAAGAAAGCTGTCTTACAAAGTTGAGCTTTGTTAGTTTTT 2100  
QY 5464 tcatgtataatattataaatt 5523  
Db TCATGTATATATTTATTAATTTATTTTCTTTGGATATATAAATGCTTTTCATATACCTG 2160  
QY 5524 ca 5525  
Db CA 2162

## RESULT 6

HSMB02391 1138 bp mRNA PRI 18-FEB-2000  
LOCUS Homo sapiens mRNA; cDNA DKFZp434P0612 (from clone DKFZp434P0612);  
DEFINITION partial cds.  
ACCESSION AL137627  
VERSION AL137627.1 GI:6808424  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1138)  
 AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY  
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.  
 This clone (DKFZp434P0612) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://www.mips.biochem.mpg.de/proj/cDNA/.

FEATURES  
 source  
 1..1138  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp434P0612"  
 /clone\_lib="434 (synonym: htes3). Vector pSport1; host DH10B; sites NotI + SalI"  
 /dev\_stage="adult"  
 /tissue\_type="testis"  
 1..576  
 /gene="DKFZp434P0612"  
 <1..576  
 /gene="DKFZp434P0612"  
 /note="Similarity to C-terminus of Drosophila eIF-2alpha kinase"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CAB70849.1"  
 /db\_xref="GI:6808425"  
 /translation="GSHVKVKSFEKEREKRVLELVDFVHVLQKLTKVTDERNGRE ASDNLAYONLKGSPVNASGLFELHIGATVPVIVSLAPEKLSASTRRYETQVQTRLQT SLANLHOKSSEILELAVDLPEKILQFLSLEWDADEQAFNTTVKLLSLRPLKQRYLKL VCEIYNIKVKVSVLLFLYSRDDYRIILF"  
 polyA\_signal 1066..1071  
 polyA\_site 1078  
 BASE COUNT 410 a 207 c 242 g 279 t  
 ORIGIN

Query Match 18.7%; Score 1032; DB 93; Length 1138;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1082; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4437 ggaagccatgtcaagggttaagctcttcgaagaagaaaggcagacagagaagcggtgtgctg 4496  
 Db 1 GGAAGCCATGTCAAGGTTAAGTCTTTTCGAGAGAAAGGAGGAGAGCGGTGTGCTG 60

Qy 4497 gagactgaactgtggaccatgtactcagaactgaggactaaagtcaactgatgaaagg 4556  
 Db 61 GAGACTGAACGTGTGGACCATGTACTCGAAGAACTGAGGACTAAGTCACTGATGAAGG 120

Qy 4557 aatggcagaagcttcgcataatctgcagtgcaaatctgaagggttcattttcatt 4616  
 Db 121 AATGGCAGAGAAGCTTCGATAATCTTCGAGTGCAAAATCTGAAGGGGTCAATTTCTAAT 180

Qy 4617 gcttcagggtttgtgaaatccatggagcaacagtgttccattgtgagtgctagcc 4676  
 Db 181 GCTTCAGGTTGTTGAAATCCATGGAGCAACAGTGGTTCCTCCATGTGAGTGGCTAGCC 240

Qy 4677 ccggagaagctgtcagccagcactagagggcgctatgaaactcaggtacaaactcgactt 4736  
 Db 241 CCGGAGAAGCTGTGAGGCGCTAGGAGCGCTATGAAGTCACTGAGTACAACTCGACTT 300

Qy 4737 cagactcccttgcgaacttacatcagaaaagcagtgaaaatttgaaattctggtgtggt 4796  
 Db 301 CAGACTCCCTTGCCAACTTACATCAGAAAAGCAGTGAATTTGAAATTTCTGGCTGGAT 360

Qy 4797 ctacccaaagaacaaatcacagttttttatcatgagtgagtgatgagtgatgagtgagtg 4856

Db 361 CTACCCAAAGAAACAATATTACAGTTTTTATCATTTAGATGGGATGTGATGAACAGCA 420  
 Qy 4857 tttaacaaactgtgaagcagctgtctgcagcctgcacagcctccaaagcaaatcctcaaat 4916  
 Db 421 TTTAAACAACACTGTGAAGCAGCTGCTGTACAGCTGCGCAAGCAAGATACCTCAAAATTA 480  
 Qy 4917 gtctgtgatgaattttataacatacaaaagtgaaaaaagtgctgtctgtctattctgtac 4976  
 Db 481 GTCTGTGATGAATTTTATAACATCAAAAGTAGAAAAAAGGTGCTGTGCTGCTATTCTGTAC 540  
 Qy 4977 agctatagagatgactactacagaactctatttttaacccctaaagaactgctgtaacctc 5036  
 Db 541 AGCTATAGATGATGACTACTACAGAACTTTATTTTAAACCCCTAAAGAAAGTGTGTTAACCTC 600  
 Qy 5037 attcaaacagacagaggtcttatactggaataatggaatgtgtacatcatcatat 5096  
 Db 601 ATTCAAAACAGACAGAGGCTTTATCTGGAATAATGGAATGTGTACATTCATCATATTTA 660  
 Qy 5097 aaattaaattcgaagagagctgggtgcagtgagtggtcacaaccttttaaccagaccttg 5156  
 Db 661 AATTAATTTCTAAGAAAGAGGCTGGGTGCAGTGGCTCACACCTTTTAAATCCACGACTTTG 720  
 Qy 5157 ggaagccaagcgaggaagactgcttgaaaccagaggttttgagaccagcctgagcaacaaa 5216  
 Db 721 GGAAGCCAAAGCAGGAGAGACTGCTTGAACCCAGAGAGTTTGAGACACAGCTTGACACAA 780  
 Qy 5217 gcaagaccccatctctataaaaaactaaaaaatttagttgggcatggtggcactgacctgt 5276  
 Db 781 GCAAGACCCCATCTCTATAAAAACTAAAAAATTAGTTGGGCATGTGTCACATGCCTGT 840  
 Qy 5277 agtccagactactccagagctgagatgagatcctgagcctcaggaggttgaggtgca 5336  
 Db 841 AGTCCAGCTACTCTCCAGAGCTGAGATGATCATCTGAGCCTCAGAGGTTTGAGGCTGCA 900  
 Qy 5337 gtgagctgtgactgcgccactgcactccagctggtggcaacagagcaagacctgtctta 5396  
 Db 901 GTCAGCTGTGACTGCGCCACTGCACCTCCAGTCTGGGCAACAGAGCAAGACCTGTCTTA 960  
 Qy 5397 aaaaaaaagaaaaaaatttttttctaaagaagctgctctacaaaagtgtgagctttgt 5456  
 Db 961 AAAAAAAGAAAAAATTTTCTTAAGAGAGCTGCTCTACAAAGTGTGAGCTTTGT 1020  
 Qy 5457 tagttttcatgtgtaataatattataatctttttgagataataataatgctttcat 5516  
 Db 1021 TAGTTTTTATGTGTAAT 1080  
 Qy 5517 ata 5519  
 Db 1081 ATA 1083

RESULT 7  
 HSM802022 1917 bp mRNA PRI 18-FEB-2000  
 LOCUS Homo sapiens mRNA; cDNA DKFZp434H149 (from clone DKFZp434H149).  
 DEFINITION AL137676  
 ACCESSION AL137676  
 VERSION AL137676.1 GI:6807846  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1917)  
 AUTHORS Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY  
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKFZp344H149) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clonerzpd.de further  
information about the clone and the sequencing project is available  
at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES		Location/Qualifiers	
Source	1..1917	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="DKFZp344H149"	
		/clone_lib="434 (synonym: htes3). Vector pspori1; host	
		DH10B; sites NotI + SalI"	
		/dev_stage="adult"	
		/tissue_type="testis"	
		1891..1896	
polyA_signal			
polyA_site			
BASE COUNT	612 a 376 c 416 g 513 t		
ORIGIN			
Query Match 16.68; Score 919; DB 93; Length 1917;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1019; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	3431	aagatattgcaagaataataattgaatttaaaacgatactgcataagaacgtgttt	3490
Db	1	AAGATATGTGGCAAGAATATATATTAATTAACGATACGCATAGAACGTGTGT	60
QY	3491	caggccgcgaagttagctgatttcacccaaagaactcttgagtggtgattgtat	3550
Db	61	CAGGCCGCGCAAGTTAGATCGATTTTCATCCCAAGAACTTCTGAGGTGTGCA	120
QY	3551	tgtcaacttaccacacacagcttctccacagctgctgaattatcacactctatga	3610
Db	121	TGTCACCTTACACCAACAGCTTTCTGCCACTGCTGAAATTAATACACTATCTA	180
QY	3611	aatcatccaagagttccagcacttcaggaagaataatcacagtattttgaaccatac	3670
Db	181	ATATCCCAAGAGTTCCAGCATTCCAGAAAGAAATTAACAGTATTATTATTA	240
QY	3671	catgtattgaagaacaactcttaccctgtggatcccaagaataaactcaagtcagt	3730
Db	241	CATGTATTGAAGCAATACCTTTACCTGTGGGATCCAGAGATAAATCACTCA	300
QY	3731	ctacattattctgtatgctgtgaacagaagctgacgagagagaagtggaagctaa	3790
Db	301	CTACATTAATCTGTATGATGCTGTGACAGAGAGCTGACGAGGAGAGAGTGGA	360
QY	3791	atttgttaactgtcttctgtcttcttaatagtctgtgtgactctacaaagtttatga	3850
Db	361	ATTGTGTAATCTGCTTTGCTTTCTTAATAGTCTGTGCTGACTCTTACAAGT	420
QY	3851	gaaggagagattgcgaagatcttatcccaacaataaattcattataaaacgaaga	3910
Db	421	GAAGGAGAGATTGCAAGATCTTATGCCAACATAAATTCATTATAAACAAGAA	480
QY	3911	tattcacagttggtgaagatgctttaaagacactagagaggtgtgtgactgttga	3970
Db	481	TATTGCACAGTTGGTGAAGATATGGCTTAAAGACCTAGAGAGGTGTGTGAC	540
QY	3971	gaactcggcatcaagttacaggtcttatcaatttgggttgggtttacaaggtgcag	4030
Db	541	GAAACTCTGCATCAAGTTACAGTCTTGATCAATTTGGGCTTGGTTTACAAG	600
QY	4031	gcacaatgaatacatcttccagttgtggtcttcttcaaaacgaagcgaaggtgtac	4090
Db	601	GCACAATGGAATCATCTTCCAGTTTGTGCTTTCTATCAAAACGAAGCGGTGT	660
QY	4091	tgaatccctgcagctggagggcagatgatgacctgctgattccccctgttagaggcc	4150
Db	661	TGAATTCCTCGAGCTGGAGGAGATATGACCTGCTGATTCCTCCAGTTTAGA	720
QY	4151	agctctggggccagttccccactgcccattgggggtcagcatagctatagacaaga	4210

Db	721	AGCTCTGGGGCCAGTTGCCACTTGGGTCAGCAGTAGCTATAGACAAGATATCTGC	780
QY	4211	tactgtcctcaacatggaggaatctgttacataaagctcttctgtgacctcctggttga	4270
Db	781	TGCTGTCTCTCAACATGGAGGAATCTGTACATAAAGCTCTTGTGTGACCTCTCTG	840
QY	4271	tgttgtcagatgtctatgtccagggccatcaacccagagaactctdgacagcagg	4330
Db	841	TGTTGGCAGATGTCTATGTCCAGGGCCATCAACCTTAACCCAGAAACTCTGGAC	900
QY	4331	catcacagcagaaatcatgtacagactggtcacagtcacagaggaattacagaagta	4390
Db	901	CATCACAGCAGAAATCATGTACGACTGTGTACAGAGGAGGAAATACAAAGATCTG	960
QY	4391	cagacatcatgaatcacctatgtggtcccttctctcgataaaggaagccatgtcaa	4450
Db	961	CAGACATCATGAATCACCTATGTGGCCCTTGTCTCGGATAAAGAGGAGGCATG	1020
QY	4451	g 4451	
Db	1021	G 1021	
RESULT 8			
AC012138/c			
LOCUS			
DEFINITION			
Homo sapiens clone RP11-10K15, WORKING DRAFT SEQUENCE, 34 unorder			
AC012138 164297 bp DNA HTG			
AC012138			
AC012138.4 GI:8954133			
HTG; HTGS_PHASE1; HTGS_DRAFT.			
SOURCE			
human.			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 164297)			
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,			
Baldwin,J., Brown,A., Castle,A., Collins,S., Collins,S., Collymore,A.,			
Cooke,P., DeArnellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,			
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,			
Gallagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,			
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,			
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,			
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Melarim,J.,			
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,			
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,			
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,			
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,			
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.			
Direct Submission			
TITLE			
JOURNAL			
COMMENT			
Submitted (20-OCT-1999) Whitehead Institute/MIT Center for Genome			
Research, 320 Charles Street, Cambridge, MA 02141, USA			
On Jul 7, 2000 this sequence version replaced by:7230052.			
All repeats were identified using RepeatMasker:			
Smit, A.F.A. & Green, P. (1996-1997)			
http://ftp.genome.washington.edu/RM/RepeatMasker.html			
----- Genome Center			
Center: Whitehead Institute/ MIT Center for Genome Research			
Center code: WIBR			
Web site: <a href="http://www-seq.wi.mit.edu">http://www-seq.wi.mit.edu</a>			
Contact: <a href="mailto:sequence_submissions@genome.wi.mit.edu">sequence_submissions@genome.wi.mit.edu</a>			
----- Project Information			
Center project name: L3086			
Center clone name: 10_K_15			
----- Summary Statistics			
Sequencing vector: M13; M77815; 98% of reads			
Sequencing vector: Plasmid; n/a; 0.0.f% of reads			



1	1064:	contig of 1064 bp in length
1065	1164:	gap of 100 bp
1165	2366:	contig of 1202 bp in length
2367	2466:	gap of 100 bp
2467	3540:	contig of 1074 bp in length
3541	3640:	gap of 100 bp
3641	4861:	contig of 1221 bp in length
4862	4961:	gap of 100 bp
4962	5968:	contig of 1007 bp in length
5969	6068:	gap of 100 bp
6069	7242:	contig of 1174 bp in length
7243	7342:	gap of 100 bp
7343	8453:	contig of 1111 bp in length
8454	8553:	gap of 100 bp
8554	9820:	contig of 1267 bp in length
9821	9920:	gap of 100 bp
9921	10938:	contig of 1018 bp in length
10939	11038:	gap of 100 bp
11039	12256:	contig of 1218 bp in length
12257	12356:	gap of 100 bp
12357	14248:	contig of 1892 bp in length
14249	14348:	gap of 100 bp
14349	15680:	contig of 1332 bp in length
15681	15780:	gap of 100 bp
15781	17240:	contig of 1460 bp in length
17241	17340:	gap of 100 bp
17341	18693:	contig of 1353 bp in length
18694	18793:	gap of 100 bp
18794	20812:	contig of 2019 bp in length
20813	20912:	gap of 100 bp
20913	24306:	contig of 3994 bp in length
24307	25006:	gap of 100 bp
25007	28926:	contig of 3920 bp in length
28927	29026:	gap of 100 bp
29027	34159:	contig of 5133 bp in length
34160	34259:	gap of 100 bp
34260	37975:	contig of 3716 bp in length
37976	38075:	gap of 100 bp
38076	43955:	contig of 5880 bp in length
43956	44055:	gap of 100 bp
44056	48039:	contig of 3984 bp in length
48040	48139:	gap of 100 bp
48140	53767:	contig of 5628 bp in length
53768	53867:	gap of 100 bp
53868	59812:	contig of 5945 bp in length
59813	59912:	gap of 100 bp
59913	63066:	contig of 3154 bp in length
63067	63166:	gap of 100 bp
63167	70627:	contig of 7461 bp in length
70628	70727:	gap of 100 bp
70728	77701:	contig of 6574 bp in length
77702	77801:	gap of 100 bp
77802	85553:	contig of 7752 bp in length
85554	85653:	gap of 100 bp
85654	94602:	contig of 8949 bp in length

Mon Oct 22 08:29:29 2001

JOURNAL  
COMMENT

Submitted (07-MAR-2000) Multimegabase Sequencing Center, University of Washington, PO BOX 357730, Seattle, WA 98195, USA  
On Apr 13, 2001 this sequence version replaced gi:10305066.

Center: Multimegabase Sequencing Center  
Genome Center  
Center code: UWMSC  
Web site: [http://chroma.mbt.washington.edu/msg\\_www](http://chroma.mbt.washington.edu/msg_www)  
Contact: leeroewensystemsbiology.org

----- Summary Statistics -----  
Sequencing vector: pUC18; L08752  
Chemistry: Dye-terminator Big Dye; 90% of reads  
Chemistry: Dye-primer Big Dye; 10% of reads  
Assembly program: Phrap; version 0.990399  
Insert size: 172000; agarose-fp  
Quality coverage: 12.6x in Q20 bases; sum-of-contigs

## Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

-----  
\* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 173364: contig of 173364 bp in length.

FEATURES  
source

Location/Qualifiers  
1. 173364  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15q14"  
/clone="RP11-43D14"  
/clone\_lib="RPCI human BAC library 11"  
/notes="This clone overlaps RP11-325N19 and RP11-521C20"  
BASE COUNT 51234 a 35556 c 35570 g 51004 t  
ORIGIN

Query Match 9.8% Score 539; DB 69; Length 173364;  
Best Local Similarity 100.0%; Pred. No. 2.2e-292;  
Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1077 cagattcaaggaacagaaacagaaattcaactcactggttaaaattgagccatccaaatgta 1136  
Db 72963 CAGATTCAAGGAACAGAAACAGAAATTCACCTCACTGCTAAATAGCCATCCAAATGTA 72904  
QY 1137 gtacgtacccttgcaatgaatctcaagagcaagcaactccatcggtggacatttta 1196  
Db 72903 GTACGTACCTTGCAATGAATCTCAAGAGCAAGACACCTCCATCGGTGGACATTTTA 72844  
QY 1197 gggagacattagtggtctctctgtgcacacactgagccactcagggcccatccct 1256  
Db 72843 GTGGAGCACATTAGTGGGGTCTCTCTGCTGCACACCTGAGCCACTCAGGCCCATCCCT 72784  
QY 1257 gtgcacagcttcgcaggtacacagctcagctcctctgagcccttgattatctgcacagc 1316  
Db 72783 GTGCATCAGCTTCGCGAGGTACACAGCTCAGCTCCTGTGTCAGCCCTTGATTATCTGCACAGC 72724  
QY 1317 aattctgtgtgcataaaggtcctgagtcacatctaatgtctgtgtggtgagcagaagccacc 1376  
Db 72723 AATTCTGTGTGCATAAAGGTCTCTGAGTCATCTAATGTCTTGGTGGATGTCAGAGGACACC 72664

misc\_feature /note="assembly\_fragment"  
77802..85553  
misc\_feature /note="assembly\_fragment"  
85654..94602  
misc\_feature /note="assembly\_fragment"  
94703..104093  
/note="assembly\_fragment"

Query Match 9.8% Score 539; DB 62; Length 164297;  
Best Local Similarity 100.0%; Pred. No. 2.2e-292;  
Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1077 cagattcaaggaacagaaacagaaattcaactcactggttaaaattgagccatccaaatgta 1136  
Db 60669 CAGATTCAAGGAACAGAAACAGAAATTCACCTCACTGCTGTAATAATGAGCCATCCAAATGTA 60610  
QY 1137 gtacgtacccttgcaatgaatctcaagagcaagcaactccatcggtggacatttta 1196  
Db 60609 GTACGCTACCTTGCAATGAATCTCAAGAGCAAGACGACTCCATCGTGTGGACATTTTA 60550  
QY 1197 gtgagcacattagtggtgtctctctgtgcacacactgagccactcagggcccatccct 1256  
Db 60549 GTGAGCACATTAGTGGGTCTCTCTGTCTGCACACCTGAGCCACTCAGGCCCATCCCT 60490  
QY 1257 gtgcacagcttcaggtacagagctcagctcctctcagggccttgattatctgcacagc 1316  
Db 60489 GTGCATCAGCTTTCGCGAGGTACACAGCTCAGCTCCTGTCAAGGCCCTTGATTATCTGCACAGC 60430  
QY 1317 aattctgtgtgcataaaggtcctgagtcgcatctaatgtctctgtggtgagtcagaaagccacc 1376  
Db 60429 AATCTGTGTGTCATAGGTCCTGAGTGCATCTAATGTCTCTGTGGATGCAGAGGCACCC 60370  
QY 1377 gtcaagattacggactatagattcttaagcgcctcgcagacatttgcaaggaggtgtg 1436  
Db 60369 GTCAAGATTACGGACTATAGCATTCTTAAGCGCTCGCAGACATTTGCAAGGAGGATGTG 60310  
QY 1437 tttagcaaacccaggtcgttttagtgacaatgctcgtcctataaacggggaagaaa 1496  
Db 60309 TTTGAGCAAAACCGAGTTCGTTTTAGTGACANAGCTCTGCTTATAAAGGGGGAAGAAA 60250  
QY 1497 ggagatgttggcgtcttggtggtcctctgtcgtcctcgcacaggaagcaaggaatgtga 1556  
Db 60249 GCAGATGTTGGCGCTTGGCGCTTCTGCTGCTGCTCCCTCAGCCAAAGACAGGAATGTGA 60190  
QY 1557 gsgtaccctgtgacacatcccttagtgaacttaccagctacattcaagatttctaaagaa 1615  
Db 60189 GAGTACCCCTGTGACCATCCCTAGTGAATCTACCAAGTCTTCAAGATTCTTCTAAGAA 60131

RESULT 9  
AC025168/c 173364 bp DNA HTG 13-APR-2001  
LOCUS Homo sapiens chromosome 15 clone RP11-43D14 map 15q14, \*\*\*  
DEFINITION SEQUENCING IN PROGRESS \*\*\*, 1 ordered pieces.  
ACCESSION AC025168  
VERSION AC025168.6 GI:13621223  
KEYWORDS HTG; HTGS\_PHASE2; HTGS\_FULLTOP.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 173364)  
AUTHORS Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R., Pate, D. and Hood, L.  
TITLE Sequencing of human chromosome 15 D15S146-D15S117 region  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 173364)  
AUTHORS Rowen, L., Madan, A., Qin, S., Abbasi, N., Baradarani, L., Birditt, B., Bloom, S., Dors, M., Dickhoff, R., Fleetwood, P., Harrison, G., James, R., Kaur, A., Madan, A., Owen, M. P., Ratcliffe, A., Shaffer, T. and Hood, L.  
TITLE Direct Submission

```

QY 1377 gtcaagattacgagctatagcatttctaagcgcctcgcagacatttcgaaggagatgtg 1436
Db 7263 CTCAGATTACGGACTATAGCATTTCTAAGCGCTCGCAGACATTTCAAGGAGGATGTG 72604
QY 1437 tttaagcaaacccgagtgcttttagtgacaatacgtctcgtcttataaacggggaagaaa 1496
Db 72603 TTTGAGCAAAACCCGAGTTCGTTTTAGTGACAATGCTCTGCCCTTATAAACGGGGAAGAAA 72544
QY 1497 ggagatgtttggcgtcttggcctctcgtcgtcgtccctcagcaagacaggaatgtgga 1556
Db 72543 GGAGATGTTGGCGTCTGGCGCTTCTGCTGCTGCTCGCCCTCAGCAAGGACAGGAATGTGGA 72484
QY 1557 gagtaccctgtgaccatccctagtgaactaccagctgactttcaagattttctaaagaa 1615
Db 72483 GAGTACCCTGTGACCATCCCTAGTGACTTACCAGCTGACITTTCAAGATTTTCTTAAGAA 72425

RESULT 10
AC021755 157900 bp DNA HTG 21-MAR-2001
LOCUS Homo sapiens chromosome 15 clone RP11-521C20 map 15q14, WORKING
DEFINITION DRAFT SEQUENCE, 6 ordered pieces.
AC021755
AC021755.6 GI:13399357
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE human.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 157900)
Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G., Kaur,A., Madan,A.,
Nesbitt,R., Traicoff,R. and Hood,L.
Sequencing of human chromosome 15 D15S146-D15S117 region
Unpublished
2 (bases 1 to 157900)
Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
and Hood,L.
Direct Submision
Submitted (20-JAN-2000) Multimegabase Sequencing Center, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
On Mar 21, 2001 this sequence version replaced gi:8272664.
-----
Center: Multimegabase Sequencing Center
Genome Center
Web site: http://chroma.mbt.washington.edu/msg_www
Web site: leetowen@systemsbiology.org
Contact: leetowen@systemsbiology.org
-----
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399
Insert size: 150000; agarose-fp
Quality coverage: 10.2x in Q20 bases; sum-of-contigs
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.

```

```

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 67961: contig of 67961 bp in length
* 67962 68061: gap of unknown length
* 68062 76767: contig of 8706 bp in length
* 76768 76867: gap of unknown length
* 76868 79623: contig of 2756 bp in length
* 79624 90362: contig of 10639 bp in length
* 90363 90462: gap of unknown length
* 90463 93020: contig of 2558 bp in length
* 93021 93120: gap of unknown length
* 93121 157900: contig of 64780 bp in length.
FEATURES
Location/Qualifiers
1..157900
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q14"
/clone="RP11-521C20"
/clone_lib="RPCI human BAC library 11"
/note="This clone overlaps RP11-43D14 and CTD-2006D8"
BASE COUNT 43137 a 35607 c 36148 g 42500 t 508 others
ORIGIN
Query Match 9.4%; Score 519; DB 66; Length 157900;
Best Local Similarity 99.8%; Pred. No. 4.4e-281;
Matches 569; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4955 ggtgtcgtgctatttctgtacagctatagagctactactacagaatcttatttaacc 5014
Db 52726 GGTGTCTGTGCTATTTCTGTACAGCTATAGAGTACTACTACAGAACTCTTATTTAACC 52785
QY 5015 cttaagaaactgcgtttaacctattcaaacagacagagcgttactactgaataatgggaat 5074
Db 52786 CTAAAGAACTGCTGTACCTACCTATTCAACACAGACAGAGGCTTACTCTGATATATGGAA 52845
QY 5075 gtgtacattatcatataatataaataatcttaagaagagctgggtcagtggtctca 5134
Db 52846 GTTGTACATTTATCATATAATTTAAATTTAAATTTCTAAGAAGAGGCTGGGTGCAGTGGCTCA 52905
QY 5135 cacccttaatccagcactttgggaagccagagcaggaagactcttgaaaccaggaggtt 5194
Db 52906 CACCTTTAATCCAGCACCCTTTGGGAAGCCAGGAGGAGGACTGCTTGAAACCCAGGAGTT 52965
QY 5195 tgagaccagcctggagcaacaaagcaagccccctctctataaaactaaaaaataatagtt 5254
Db 52966 TGAGACCAGCTGAGCAACAAAGCAAGACCCCATCTCTATAAANAATAAATAATAGTT 53025
QY 5255 gggcatggtggcacatgctctagtcctcagctactccagagctgagatgagatcatctga 5314
Db 53026 GGGCATGTTGTCACATGCTCTAGTCCAGCTACTTCCAGAGGCTGAGATGATCATCTGA 53085
QY 5315 gcctcagagaggttgagcgtcagtgagcgtgactgcacactgcactcagctcagctggac 5374
Db 53086 GCCTCAGAGGTTGAGGCTGCGAGTGTGACTGTGCCACTGCCTCAGCTCAGCTGGGAC 53145
QY 5375 aacagagcaagaccctgtcttaaaaaaagaaaaaataatctttcttaagaagct 5434
Db 53146 AACAGAGCAAGACCCTGTCTTAAAAAAGAAAAAAGAAAAAATTTTCTTAAGAGACT 53205
QY 5435 gtccatacaaaattgagctttgttagtttttcattgtgtaataatatataaatttatctttt 5494
Db 53206 GTCCTACAAAGTTGAGCTTTGTGTTTTCATGTTGTAATATATATAAATTTATCTTTT 53265
QY 5495 gggataataaattgctttctatatacctgc 5524
Db 53266 GGGATATAATAAATGCTTTTATATACCTGC 53295
RESULT 11

```

AC090997 61029 bp DNA HTG 22-MAR-2001  
LOCUS Homo sapiens chromosome 15 clone RP11-521C20 map 15, LOW-PASS  
DEFINITION SEQUENCE SAMPLING.  
AC090997  
VERSION AC090997.1 GI:13431057  
KEYWORDS HTG; HTGS\_PHASE0.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 61029)  
Birten,B., Linton,L., Nusbaum,C. and Lander,E.  
TITLE Homo sapiens chromosome 15, clone RP11-521C20  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 61029)  
Birten,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,  
Barne,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,  
Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,  
Collins,S., Collymore,A., Cooke,P., DeArelano,K., Dewar,K., D.,  
Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,  
Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K.,  
Lamarez,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
Maclean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,  
McKernan,P., McKernan,K., McPheters,R., Meldrum,J., Meneus,L.,  
Milova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,  
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,  
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,  
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,  
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,  
Severy,P., Sougez,C., Spencer,B., Stange-Thomann,N.,  
Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S.,  
Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (22-MAR-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L13141  
Center clone name: 521\_C-20  
-----  
\* NOTE: This record contains 82 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
\* 1  
\* 571 670: gap of 100 bp  
\* 671 1297: contig of 627 bp in length  
\* 1298 1397: gap of 100 bp  
\* 1398 2087: contig of 690 bp in length  
\* 2088 2187: gap of 100 bp  
\* 2188 2956: contig of 769 bp in length  
\* 2957 3056: gap of 100 bp  
\* 3057 3690: contig of 634 bp in length  
\* 3691 3790: gap of 100 bp  
\* 3791 4405: contig of 615 bp in length  
\* 4406 4505: gap of 100 bp  
\* 4506 5161: contig of 656 bp in length  
\* 5162 5261: gap of 100 bp  
\* 5262 5867: contig of 606 bp in length  
\* 5868 5967: gap of 100 bp  
\* 5968 6630: contig of 663 bp in length  
\* 6631 6730: gap of 100 bp  
\* 6731 7361: contig of 631 bp in length  
\* 7362 7461: gap of 100 bp  
\* 7462 8083: contig of 622 bp in length  
\* 8084 8183: gap of 100 bp  
\* 8184 8853: contig of 670 bp in length  
\* 8854 8953: gap of 100 bp  
\* 8954 9677: contig of 724 bp in length  
\* 9678 9777: gap of 100 bp  
\* 9778 10498: contig of 721 bp in length  
\* 10499 10598: gap of 100 bp  
\* 10599 11211: contig of 613 bp in length  
\* 11212 11311: gap of 100 bp  
\* 11312 11946: contig of 635 bp in length  
\* 11947 12046: gap of 100 bp  
\* 12047 12683: contig of 636 bp in length  
\* 12683 12782: gap of 100 bp  
\* 12783 13443: contig of 661 bp in length  
\* 13444 13543: gap of 100 bp  
\* 13544 14212: contig of 669 bp in length  
\* 14213 14312: gap of 100 bp  
\* 14313 14919: contig of 607 bp in length  
\* 14920 15019: gap of 100 bp  
\* 15020 15680: contig of 661 bp in length  
\* 15681 15780: gap of 100 bp  
\* 15781 16337: contig of 557 bp in length  
\* 16338 16437: gap of 100 bp  
\* 16438 17112: contig of 675 bp in length  
\* 17113 17212: gap of 100 bp  
\* 17213 17901: contig of 689 bp in length  
\* 17902 18001: gap of 100 bp  
\* 18002 18749: contig of 748 bp in length  
\* 18750 18849: gap of 100 bp  
\* 18850 19453: contig of 604 bp in length  
\* 19454 19553: gap of 100 bp  
\* 19554 20189: contig of 636 bp in length  
\* 20190 20289: gap of 100 bp  
\* 20290 20942: contig of 653 bp in length  
\* 20943 21042: gap of 100 bp  
\* 21043 21648: contig of 606 bp in length  
\* 21649 21748: gap of 100 bp  
\* 21749 22357: contig of 609 bp in length  
\* 22358 22457: gap of 100 bp  
\* 22458 23104: contig of 647 bp in length  
\* 23105 23204: gap of 100 bp  
\* 23205 23840: contig of 636 bp in length  
\* 23841 23940: gap of 100 bp  
\* 23941 24516: contig of 576 bp in length  
\* 24517 24616: gap of 100 bp  
\* 24617 25314: contig of 698 bp in length  
\* 25315 25414: gap of 100 bp  
\* 25415 26037: contig of 623 bp in length  
\* 26038 26137: gap of 100 bp  
\* 26138 26749: contig of 612 bp in length  
\* 26750 26849: gap of 100 bp  
\* 26850 27497: contig of 648 bp in length  
\* 27498 27597: gap of 100 bp  
\* 27598 28259: contig of 662 bp in length  
\* 28260 28359: gap of 100 bp  
\* 28360 29041: contig of 682 bp in length  
\* 29042 29141: gap of 100 bp  
\* 29142 29807: contig of 666 bp in length  
\* 29808 29907: gap of 100 bp  
\* 29908 30480: contig of 573 bp in length  
\* 30481 30580: gap of 100 bp  
\* 30581 31289: contig of 709 bp in length

Db	48163	AGTGAGCTGTGACTGCGCCACTGCACCTCCAGTCTGGGACACAGACGACCCCTGTCTT	48222
QY	5396	aaaaaaaaaagaaaaaaatTTTTTTTctaaagaagctgtctcacaagttagctttg	5455
Db	48223	AAAAAAAAAAGAAAAAATTTTTTCTTAAGAAGCTGCTCTTACAAGTTGAGCTTTG	48282
QY	5456	ttagtttttcattg	5468
Db	48283	TTAGTTTTTTCATG	48295
RESULT	12		
AC090997/c			
LOCUS			
DEFINITION		AC090997 61029 bp DNA HTG 22-MAR-2001	
ACCESSION		Homo sapiens chromosome 15 clone RP11-521C20 map 15, LOW-PASS	
VERSION		SEQUENCE SAMPLING.	
KEYWORDS		AC090997.1 GI:13431057	
SOURCE		HTG; HTGS_PHASE0.	
ORGANISM		human.	
REFERENCE		Homo sapiens	
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE		1 (bases 1 to 61029)	
JOURNAL		Birren,B., Linton,L., Nusbaum,C. and Lander,E.	
AUTHORS		Homo sapiens chromosome 15, clone RP11-521C20	
		Unpublished	
		2 (bases 1 to 61029)	
		Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barina,N., Bastien,V., Boguslavsky,L., Bouckhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severyanov,N., Strauss,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
TITLE		Direct Submission	
JOURNAL		Submitted (22-MAR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
COMMENT		All repeats were identified using RepeatMasker: Smit,A.F.A. & Green, P. (1996-1997) <a href="http://ftp.genome.washington.edu/RM/RepeatMasker.html">http://ftp.genome.washington.edu/RM/RepeatMasker.html</a>	
		----- Genome Center	
		Center: Whitehead Institute/ MIT Center for Genome Research	
		Center code: WIBR	
		Web site: <a href="http://www-seq.wi.mit.edu">http://www-seq.wi.mit.edu</a>	
		Contact: <a href="mailto:sequence_submissions@genome.wi.mit.edu">sequence_submissions@genome.wi.mit.edu</a>	
		----- Project Information	
		Center project name: L1341	
		Center clone name: 521_C_20	
		-----	
		* NOTE: This record contains 82 individual	
		* sequencing reads that have not been assembled into	
		* contigs. Runs of N are used to separate the reads	
		* and the order in which they appear is completely	
		* arbitrary. Low-pass sequence sampling is useful for	
		* identifying clones that may be gene-rich and allows	
		* overlap relationships among clones to be deduced.	
		* However, it should not be assumed that this clone	

\* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

```

1
571 670: contig of 570 bp in length
671 1297: contig of 627 bp in length
1298 1397: gap of 100 bp
1398 2087: contig of 690 bp in length
2088 2187: gap of 100 bp
2188 2956: contig of 769 bp in length
2957 3056: gap of 100 bp
3057 3650: contig of 634 bp in length
3651 3790: gap of 100 bp
3791 4405: contig of 615 bp in length
4406 4505: gap of 100 bp
4506 5161: contig of 656 bp in length
5162 5261: gap of 100 bp
5262 5867: contig of 606 bp in length
5868 5967: gap of 100 bp
5968 6630: contig of 663 bp in length
6631 6730: gap of 100 bp
6731 7361: contig of 631 bp in length
7362 7461: gap of 100 bp
7462 8083: contig of 622 bp in length
8084 8183: gap of 100 bp
8184 8853: contig of 670 bp in length
8854 8953: gap of 100 bp
8954 9677: contig of 724 bp in length
9678 9777: gap of 100 bp
9778 10498: contig of 721 bp in length
10499 10598: gap of 100 bp
10599 11211: contig of 613 bp in length
11212 11311: gap of 100 bp
11312 11946: contig of 635 bp in length
11947 12046: gap of 100 bp
12047 12682: contig of 636 bp in length
12683 12782: gap of 100 bp
12783 13443: contig of 661 bp in length
13444 13543: gap of 100 bp
13544 14212: contig of 669 bp in length
14213 14312: gap of 100 bp
14313 14919: contig of 607 bp in length
14920 15019: gap of 100 bp
15020 15680: contig of 661 bp in length
15681 15780: gap of 100 bp
15781 16337: contig of 557 bp in length
16338 16437: gap of 100 bp
16438 17112: contig of 675 bp in length
17113 17212: gap of 100 bp
17213 17901: contig of 689 bp in length
17902 18001: gap of 100 bp
18002 18749: contig of 748 bp in length
18750 18849: gap of 100 bp
18850 19453: contig of 604 bp in length
19454 19553: gap of 100 bp
19554 20189: contig of 636 bp in length
20190 20289: gap of 100 bp
20290 20943: contig of 653 bp in length
20943 21042: gap of 100 bp
21043 21648: contig of 606 bp in length
21649 21748: gap of 100 bp
21749 22357: contig of 609 bp in length
22358 22457: gap of 100 bp
22458 23104: contig of 647 bp in length
23105 23204: gap of 100 bp
23205 23840: contig of 636 bp in length
23841 23940: gap of 100 bp
23941 24516: contig of 576 bp in length
24517 24616: gap of 100 bp
24617 25314: contig of 698 bp in length
25315 25414: gap of 100 bp
25415 26037: contig of 623 bp in length
26038 26137: gap of 100 bp

```

```

26138 26749: contig of 612 bp in length
26750 26849: gap of 100 bp
26850 27497: contig of 648 bp in length
27498 27597: gap of 100 bp
27598 28259: contig of 662 bp in length
28260 28359: gap of 100 bp
28360 29041: contig of 682 bp in length
29042 29141: gap of 100 bp
29142 29807: contig of 666 bp in length
29808 29907: gap of 100 bp
29908 30480: contig of 573 bp in length
30481 30580: gap of 100 bp
30581 31289: contig of 709 bp in length
31290 31389: gap of 100 bp
31390 32097: contig of 708 bp in length
32098 32197: gap of 100 bp
32198 32914: contig of 717 bp in length
32915 33014: gap of 100 bp
33015 33618: contig of 604 bp in length
33619 33718: gap of 100 bp
33719 34358: contig of 640 bp in length
34359 34458: gap of 100 bp
34459 34977: contig of 519 bp in length
34978 35077: gap of 100 bp
35078 35689: contig of 612 bp in length
35690 35789: gap of 100 bp
35790 36386: contig of 597 bp in length
36387 36486: gap of 100 bp
36487 37092: contig of 606 bp in length
37093 37192: gap of 100 bp
37193 37864: contig of 672 bp in length
37865 37964: gap of 100 bp
37965 38626: contig of 662 bp in length
38627 38726: gap of 100 bp
38727 39348: contig of 622 bp in length
39349 39448: gap of 100 bp
39449 40072: contig of 624 bp in length
40073 40172: gap of 100 bp
40173 40882: contig of 710 bp in length
40883 40982: gap of 100 bp
40983 41705: contig of 723 bp in length
41706 41805: gap of 100 bp
41806 42412: contig of 607 bp in length
42413 42512: gap of 100 bp
42513 43158: contig of 646 bp in length
43159 43258: gap of 100 bp
43259 43880: contig of 622 bp in length
43881 43980: gap of 100 bp
43981 44618: contig of 638 bp in length
44619 44718: gap of 100 bp
44719 45359: contig of 641 bp in length
45360 45459: gap of 100 bp
45460 46088: contig of 629 bp in length
46089 46188: gap of 100 bp
46189 46855: contig of 667 bp in length
46856 46955: gap of 100 bp
46956 47642: contig of 687 bp in length
47643 47742: gap of 100 bp
47743 48295: contig of 553 bp in length
48296 48395: gap of 100 bp
48396 49021: contig of 626 bp in length
49022 49121: gap of 100 bp
49122 49825: contig of 704 bp in length
49826 49925: gap of 100 bp
49926 50543: contig of 618 bp in length
50544 50643: gap of 100 bp
50644 51266: contig of 623 bp in length
51267 51366: gap of 100 bp

```

Query Match 3.1%; Score 171; DB 78; Length 61029;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-84;  
 Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```

QY 3469 gatactgataaagctgttccagcgccgaagttagatcattccatcccaagaac 3528
Db 13127 GATATGTCATAGACGTTGTTTCAGCGCGCGCAAGTTAGATGATTTCATCCCAAGAAC 13068
QY 3529 ttctggagtgcatttgattgttcaactttacacacacagctttctgcccactgctg 3588
Db 13067 TTCTGGAGTGTGCAATTTGATATTGTCACTTCTACCAACACAGCTTCTGCCACTGCTG 13008
QY 3589 aaattatcacactatcatataaatacccaagagtttccagcaacttcagg 3639
Db 13007 AATTTATCTACACTATCATGAATCATCCCAAGAGTTTCCAGCACTTCAGG 12957

RESULT 13
AC012377
LOCUS AC012377 197610 bp DNA HTG 13-APR-2001
DEFINITION Homo sapiens chromosome 15 clone RP11-325N19 map 15q14, ***
SEQUENCING IN PROGRESS ***, 1 ordered pieces.
AC012377
VERSION AC012377.4 GI:13621218
KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 197610)
Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B.,
Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
Pate,D. and Hood,L.
Sequencing of human chromosome 15 D15S146-D15S117 region
Unpublished
2 (bases 1 to 197610)
Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
Bloom,S., Dors,M., Dickhoff,R., Harrison,G., James,K., Madan,A.,
Owen,M.P., Ratcliffe,A., Shaffer,T. and Hood,L.
Direct Submission
Submitted (26-OCT-1999), Multimegabase Sequencing Center, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
On Apr 13, 2001 this sequence version replaced gi:10567837.
----- Genome Center Sequencing Center
Center: Multimegabase Sequencing Center
Center code: UMSC
Web site: http://chroma.mbt.washington.edu/msg_www
Contact: leerowensystemsbiology.org
----- Summary Statistics
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Assembly: Dye-primer Big Dye; 10% of reads
Insert size: 197000; agarose-fp
Quality coverage: 13x in Q20 bases; sum-of-contigs
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as runs of N. However the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 197610: contig of 197610 bp in length.
* Location/Qualifiers

```

## source

```

1. 197610
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q14"
/clone="RP11-325N19"
/clone_lib="human RPCI BAC library 11"
/note="This clone overlaps RP11-37C7 and RP11-43D14"
BASE COUNT 61560 a 41034 c 39114 g 55902 t
ORIGIN

```

```

Query Match 2.8%; Score 155; DB 62; Length 197610;
Best Local Similarity 100.0%; Pred. NO. 3e-75;
Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 421 agtgatgatcttgaactggcttaccagctgcagtcattctcagcagcatacaagc 480
Db 188078 AGTGATGATCTTTGAACCTGGCTTACCAGTGCAGTCATTTCTCAGCAGCATAACAAGC 188137
QY 481 cccctcccaagtccttttcatgaagaatgctggaagcgcgctcagggagcagcaga 540
Db 188138 CCCTCCCAAGTCTTTTCATGAAGAAATGCTGAAAGCGGGCTCAGGAGGAGCAGCAGA 188197
QY 541 ggctgttgagcccaagcggaaagagcagcagag 575
Db 188198 GGCTGTGGAGCGCAAGCGGAAAGAGCAGCAGAG 188232

```

## RESULT 14

```

AC012138 164297 bp DNA HTG 07-JUL-2000
LOCUS Homo sapiens clone RP11-10K15, WORKING DRAFT SEQUENCE, 34 unordered
DEFINITION pieces.
AC012138
VERSION AC012138.4 GI:8954133
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 164297)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-10K15
Unpublished
2 (bases 1 to 164297)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,
Cooke,P., Dearrellano,K., Dewar,K., Domino,M., Donegan,L., Doyle,M.,
Ferrelira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., Norman,C.H., O'Connor,T., O'Donnell,P.,
Morrow,J., Naylor,J., Naylor,R., Roy,A., Santos,R., Severy,P.,
Stange-Thoman,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tefaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye.W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 7, 2000 this sequence version replaced gi:7230052.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information

```

## FEATURES

Center project name: L3086  
Center clone name: 10\_K\_15  
----- Summary Statistics  
Sequencing vector: M13; M7815; 98% of reads  
Sequencing vector: Plasmid; n/a; %-0.f%% of reads  
1.91489361702128Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 142640 bases at least Q40  
Consensus quality: 152432 bases at least Q30  
Consensus quality: 157096 bases at least Q20  
Insert size: 153000; agarose-fp  
Insert size: 160997; sum-of-contigs  
Quality coverage: 3.4 in Q20 bases; agarose-fp  
Quality cov.

NOTE: This is a 'working draft' sequence. It currently consists of 34 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1064: contig of 1064 bp in length  
\* 1065 1164: gap of 100 bp  
\* 1165 2366: contig of 1202 bp in length  
\* 2367 2466: gap of 100 bp  
\* 2467 3540: contig of 1074 bp in length  
\* 3541 3640: gap of 100 bp  
\* 3641 4861: contig of 1221 bp in length  
\* 4862 4961: gap of 100 bp  
\* 4962 5968: contig of 1007 bp in length  
\* 5969 6068: gap of 100 bp  
\* 6069 7242: contig of 1174 bp in length  
\* 7243 7342: gap of 100 bp  
\* 7343 8453: contig of 1111 bp in length  
\* 8454 8553: gap of 100 bp  
\* 8554 9820: contig of 1267 bp in length  
\* 9821 9920: gap of 100 bp  
\* 9921 10938: contig of 1018 bp in length  
\* 10939 11038: gap of 100 bp  
\* 11039 12256: contig of 1218 bp in length  
\* 12257 12356: gap of 100 bp  
\* 12357 14248: contig of 1892 bp in length  
\* 14249 14348: gap of 100 bp  
\* 14349 15680: contig of 1332 bp in length  
\* 15681 15780: gap of 100 bp  
\* 15781 17240: contig of 1460 bp in length  
\* 17241 17340: gap of 100 bp  
\* 17341 18693: contig of 1353 bp in length  
\* 18694 18793: gap of 100 bp  
\* 18794 20812: contig of 2019 bp in length  
\* 20813 20912: gap of 100 bp  
\* 20913 24906: contig of 3994 bp in length  
\* 24907 25006: gap of 100 bp  
\* 25007 28926: contig of 3920 bp in length  
\* 28927 29026: gap of 100 bp  
\* 29027 34159: contig of 5133 bp in length  
\* 34160 34259: gap of 100 bp  
\* 34260 37975: contig of 3716 bp in length  
\* 37976 38075: gap of 100 bp  
\* 38076 43955: contig of 5880 bp in length  
\* 43956 44055: gap of 100 bp  
\* 44056 48039: contig of 3984 bp in length  
\* 48040 48139: gap of 100 bp  
\* 48140 53767: contig of 5628 bp in length  
\* 53768 53867: gap of 100 bp  
\* 53868 59812: contig of 5945 bp in length  
\* 59813 59912: gap of 100 bp  
\* 59913 63066: contig of 3154 bp in length  
\* 63067 63166: gap of 100 bp  
\* 63167 70627: contig of 7461 bp in length  
\* 70628 70727: gap of 100 bp

## FEATURES

source

Location/Qualifiers  
1. 164297  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-10K15"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. 1064  
/note="assembly\_fragment"  
1165. 2366  
/note="assembly\_fragment"  
2467. 3540  
/note="assembly\_fragment"  
3641. 4861  
/note="assembly\_fragment"  
4962. 5968  
/note="assembly\_fragment"  
6069. 7242  
/note="assembly\_fragment"  
7343. 8453  
/note="assembly\_fragment"  
8554. 9820  
/note="assembly\_fragment"  
9921. 10938  
/note="assembly\_fragment"  
11039. 12256  
/note="assembly\_fragment"  
12357. 14248  
/note="assembly\_fragment"  
14349. 15680  
/note="assembly\_fragment"  
15781. 17240  
/note="assembly\_fragment"  
17341. 18693  
/note="assembly\_fragment"  
18794. 20812  
/note="assembly\_fragment"  
20913. 24906  
/note="assembly\_fragment"  
25007. 28926  
/note="assembly\_fragment"  
29027. 34159  
/note="assembly\_fragment"  
34260. 37975  
/note="assembly\_fragment"  
clone\_end:17  
vector\_side:left  
38076. 43955  
/note="assembly\_fragment"  
44056. 48039  
/note="assembly\_fragment"  
48140. 53767  
/note="assembly\_fragment"  
53868. 59812  
/note="assembly\_fragment"  
59913. 63066  
/note="assembly\_fragment"  
/note="assembly\_fragment"

\* 70728 77701: contig of 6974 bp in length  
\* 77702 77801: gap of 100 bp  
\* 77802 85553: contig of 7752 bp in length  
\* 85554 85653: gap of 100 bp  
\* 85654 94602: contig of 8949 bp in length  
\* 94603 94702: gap of 100 bp  
\* 94703 104093: contig of 9391 bp in length  
\* 104094 104193: gap of 100 bp  
\* 104194 112395: contig of 8202 bp in length  
\* 112396 112495: gap of 100 bp  
\* 112496 123244: contig of 10749 bp in length  
\* 123245 123344: gap of 100 bp  
\* 123345 135427: contig of 12083 bp in length  
\* 135428 135527: gap of 100 bp  
\* 135528 147434: contig of 11907 bp in length  
\* 147435 147534: gap of 100 bp  
\* 147535 164297: contig of 16763 bp in length.



